WebSci'18
Linked Learning 2018

27 May 2018, VU Amsterdam, The Netherlands

7th International Workshop on Learning and Education with Web Data (#LILE2018)

Chairs & Editors:

Stefan Dietze, L3S Research Center, Germany
Mathieu D'Aquin, Insight Centre for Data Analytics, Ireland
Dragan Gasevic, Monash University, Australia
Eelco Herder, L3S Research Center, Germany
Joachim Kimmerle, Knowledge Media Research Center, Germany

Linked Learning 2018 - Learning and Education with Web Data

Distance teaching and the use of openly available educational resources on the Web are becoming common practices at public higher education institutions as well as private training organisations. In addition, informal learning and knowledge exchange are inherent to our daily online interactions, such as searching the Web [1], and using learning and knowledge-centric social networks, such as Bibsonomy, Slideshare, Wikipedia and Videolectures, or general-purpose social environments such as LinkedIn, where matters related to skills, competence development or training are central concerns of involved stakeholders. These interactions generate a vast amount of data, about informal knowledge resources of varying granularity as well as user activities, including informal indicators for learning and competences.

At the same time, the prevalence of entity-centric Web data, facilitated through Open Data, Knowledge Graphs or Linked Data [2], as well as the more recent widespread adoption of embedded annotations through schema.org, Microdata and RDFa has led to the availability of vast amounts of semi-structured data which facilitates interpretation and reuse of Web content and data in learning scenarios [3].

Initiatives such as LinkedUp or the more recent AFEL project[1] have already made available collections of learning-related data, covering both user activity and resource-centric information. The widespread analysis of both informal and formal learning activities and resources has the potential to fundamentally aid and transform the production, recommendation and consumption of learning services and content. Typical scenarios include the use of machine learning for automatically classifying learning performance, competences or user knowledge, by learning from the vast amounts of available data or, to exploit resource-centric data and knowledge graphs to automatically generate learning resources or assessment items.

However, interpreting learning activities and online interactions requires a highly interdisciplinary skillset, including knowledge about learning theory, psychology and sociology, but also technical means to enable data analysis in large-scale heterogeneous data. Building on the success of several editions, LILE2018 addresses such challenges by providing a forum for researchers and practitioners who make innovative use of Web data for educational purposes, spanning areas such as learning analytics, Web mining, data and Web science, psychology and the social sciences.

After extensive peer review (each submission was reviewed by at least three independent reviewers) we have been able to select six papers for presentation in the program. LILE2018 also featured two keynotes, addressing learning-related topics from both technical as well social sciences perspectives.

The workshop would not have been possible without contributions of many people and institutions. We are very thankful to the organizers of the ACM Web Science 2018 conference for providing us with the opportunity to organize the workshop, for their excellent collaboration, and for looking after many logistic issues. We are also very grateful to the members of the program committee for their commitment in reviewing the papers and assuring the good quality of the

[1] http://afel-project.eu

workshop program. We also thank all authors and most importantly, our keynote speakers John Domingue (The Open University, UK) and Inge Molenaar (Radboud University, NL) for their invaluable contributions to the workshop. Finally, great appreciation goes to our sponsors GNOSS[2] and AFEL for their support.

REFERENCES

[1] Yu, R., Gadiraju, U., Holtz, P., Rokicki, M., Kemkes, P., Dietze, S., Predicting User Knowledge Gain in Informational Search Sessions, full research track paper at 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR2018), Ann Arbor Michigan, U.S.A. July 8-12, 2018, ACM.

[2] D'Aquin, M., Adamou, A., Dietze, S. 2013. Assessing the Educational Linked Data Landscape. *In Proceedings of ACM Web Science 2013 (WebSci2013)*, Paris, France, May 2013.

[3] Dietze, S., Taibi, D., Yu, R., Barker, P., d'Aquin, M., Analysing and Improving embedded Markup of Learning Resources on the Web, 26th International World Wide Web Conference (WWW2017), Digital Learning track, Perth, Australia, April 2017.

LINKED LEARNING 2018 – ACCEPTED PAPERS

- Seren Yenikent, Brett Buttliere, Besnik Fetahu and Joachim Kimmerle. *Wikipedia Article Measures in relation to Content Characteristics of Lead Sections*
- Tatiana Person, Iván Ruiz-Rube and Juan Manuel Dodero. *Exploiting the Web of Data for the creation of mobile apps by non-expert programmers*
- Simone Kopeinik, Almonzer Eskandar, Tobias Ley, Dietrich Albert and Paul Seitlinger. *Adapting an open source social bookmarking system to observe critical information behaviour*
- Anett Hoppe, Peter Holtz, Yvonne Kammerer, Ran Yu, Stefan Dietze and Ralph Ewerth. *Current Challenges for Studying Search as Learning Processes*
- Sven Lieber, Ben De Meester, Anastasia Dimou and Ruben Verborgh. *Linked Data Generation for Adaptive Learning Analytics Systems*
- Ran Yu, Ujwal Gadiraju and Stefan Dietze. *Detecting, Understanding and Supporting Everyday Learning in Web Search*

[2] http://gnoss.com/

LINKED LEARNING 2018 – ORGANIZATION

Workshop Chairs:

Stefan Dietze, L3S Research Center, Germany
Mathieu d'Aquin, Insight Centre for Data Analytics, Ireland
Dragan Gasevic, Monash University, Australia
Eelco Herder, L3S Research Center, Germany
Joachim Kimmerle, Knowledge Media Research Centre, Germany

Program Committee:

Erik Barendsen, Radboud University & Open University
Jürgen Buder, Leibniz-Institut für Wissensmedien
Brett Buttliere, IWM Tübingen
Marco Antonio Casanova, PUC – Rio
Ulrike Cress, Knowledge Media Research Center
Michel Desmarais, Ecole Polytechnique de Montreal
John Domingue, The Open University
Nikolas Dovrolis, Democritus University of Thrace
Angela Fessl, Know-Center, Austria
Christophe Guéret, Accenture
Claudia Hauff, Delft University of Technology
Maurice Hendrix, Coventry University
Peter Holtz, Leibniz Insitut für Wissensmedien Tübingen
Jelena Jovanovic, University of Belgrade
Carsten Keßler, Department of Planning, Aalborg University Copenhagen
Elisabeth Lex, Graz University of Technology
Johannes Moskaliuk, International School of Management
Dmitry Mouromtsev, NRU ITMO, Russia
Bernardo Pereira Nunes, PUC-Rio
Niels Pinkwart, Humboldt-Universität zu Berlin
Carolyn Rose, Carnegie Mellon University
Sergey Sosnovsky, Utrecht University
Mari Carmen Suárez-Figueroa, Universidad Politécnica de Madrid
Nadine Steinmetz, TU Ilmenau
Davide Taibi, Italian National Research Council
Fridolin Wild, Oxford Brookes University
Ran Yu, L3S Research Center

Wikipedia Article Measures in relation to Content Characteristics of Lead Sections

Seren Yenikent
Leibniz-Institut für Wissensmedien
Tuebingen, Germany
s.yenikent@iwm-tuebingen.de

Brett Buttliere
Technical University Dresden
Dresden, Germany
brettbuttliere@gmail.com

Besnik Fetahu
L3S Research Center
Hannover, Germany
fetahu@L3S.de

Joachim Kimmerle
Leibniz-Institut für Wissensmedien
Tuebingen, Germany
j.kimmerle@iwm-tuebingen.de

ABSTRACT

In this study we examined the sentimental and psychological content in Wikipedia articles' lead sections; and studied how this content was related to the articles' descriptive measures. First, two text analysis tools were used to determine positive and negative content, which were both positively related to the article measures. A closer examination found that particularly psychological drive states (topics of achievement, reward, risk, affiliation, and power) and affective processes (positive and negative emotion) were the most related to measures like article length and number of links. These findings suggest that descriptive article characteristics are associated with the sentimental and psychological content of the lead section.

KEYWORDS

Wikipedia, textual analysis, sentiment analysis, LIWC

1 INTRODUCTION

Wikipedia content represents contemporary life by reflecting the knowledge creation and consumption of the public [15]. Considering the active role of editors in the knowledge construction process, it is important to understand what motivates them to contribute to the article content. Previous studies that assessed Wikipedia articles' content have mostly utilized computer science and data science concepts. In this study we take a psychological perspective and examine whether the descriptive features of an article, such as length, number of links, number of sections, and number of images, are related to the sentiment and psychological content of this article's lead section.

WebSci '18, May 27-30 2018, Amsterdam, Netherlands

The lead section of a Wikipedia article is the introduction and summary of the presented topic, and thus is the most representative section. Among the few studies that focused on lead sections, Wagner and colleagues [18] demonstrated that these sections can be used to better understand societal dynamics of Wikipedia. We examined the content characteristics of lead sections on two levels; (1) from a broader point of view, by assessing the *positive* and *negative* sentiment; and (2) from a more specific exploration based on the underlying *psychological processes*. Our aim was to understand how the descriptive measures of articles were related to these sentimental and psychological variables.

1.1 Positive and negative sentiment of content

People tend to define their experiences based on positive and negative emotional dimensions [2]. Mixed emotion theories suggest that positive and negative emotions affect human behavior by being simultaneously activated whereas dimensional models posit opposing ends for positive-negative emotions. In this study, we aimed to understand how these emotional aspects are reflected in the content of Wikipedia lead sections and related to article features. In order to gauge the positive and negative orientation of the words in a systematic way, we took up a sentiment analysis approach.

Sentiment analysis refers to the field that uses methods of natural language processing to automatically identify the polarity of a text – whether the text has a positive or negative orientation – in order to analyze people's opinions and attitudes toward certain entities and topics [10]. Hu and Liu [8], for instance, identified opinion words and developed a dictionary by adopting an approach that associated positive words with desirable states and negative words with undesirable states. In our study we employed a similar semantic approach in the operationalization of Wikipedia content: positive content in reference to desirable and joyful experiences, and negative content in reference to undesirable experiences, such as conflicts and controversies.

Although a major line of research highlights the role of negative content (e.g., controversies) in online environments, there is also evidence that positive content (e.g., inspiring texts) prompts interest, whereby both types may lead to larger communities with longer and richer content [1]. Former

empirical studies on the effects of positive and negative content on Wikipedia activities also found rather mixed results. Wikipedia editors were reported as adopting a positive attitude and expressing positive emotions [9]. However, negative Wikipedia content that includes, for instance, controversial topics prompt more interest and receive more edits [20, 21]. These ambiguous results make it worthwhile to scrutinize the issue of positive and negative content and article creation in Wikipedia.

1. 2 Content based on psychological processes

Words people use provide insights into their psychological experiences [13]. Although Wikipedia ensures objectivity and neutral perspectives, editors' characteristics and psychological experiences are still reflected in the articles' content. Previous studies showed that emotional and psychological responses may spill over to certain types of articles, for instance, in the forms of anger- and sadness-related or group-biased content [7, 11].

In order to examine the psychologically relevant content in the Wikipedia articles, we utilized the framework of Linguistic Inquiry and Word Count (LIWC) text analysis approach, which was specially developed for the purpose of identifying psychological processes in texts [13]. LIWC offers a rich variety of content dimensions that provides solid analysis options for psychological processes behind concepts such as thinking style, social relationships, and group processes (for a detailed overview see [17]). The psychological dimensions of LIWC have been successfully applied to Wikipedia articles. Several studies revealed the dynamics behind specific content, such as disasters and terrorist attacks, and their relation to Wikipedia activities [6, 7]. We followed a similar research idea with the aim of associating LIWC's psychological framework with the Wikipedia articles' descriptive measures. Yet, our study goes one step further and focuses on all of the psychological content instead of specific topics.

1.3 Article measures

We aggregated four different measures that are simple yet robust metrics of Wikipedia articles: *article length, number of links, number of sections*, and *number of images*. We assume that these simple metrics may characterize the nature of an article and provide information on how elaborate and comprehensive an article is [3, 19, 22].

We brought these article measures together under a single metric and related it to the Wikipedia lead sections. Considering that the process of content creation is realized through producing and/or changing text along with other features, such as images and links, we expected a certain type of relationship between content characteristics and articles' measures.

Specifically, we asked the following research questions to understand the sentiment and psychological characteristics of the lead sections and comprehend how they were associated with the measures of an article as a whole:

RQ1: Does positive and negative content differ in the Wikipedia lead sections?

RQ2: To what extent is positive and negative content in Wikipedia lead sections related to the article measures?

RQ3: To what extent are different kinds of content based on psychological processes related to the article measures?

2 METHODS

2.1 Dataset

A dataset with a sample of 63,289 articles was randomly extracted from the society portal of the English language version of Wikipedia (June 2017). The dataset contained lead sections of Wikipedia articles and article measures including article length in characters, number of links, number of sections, and number of images.

2.2 Article metric

In order to create a single metric representing article measures, a factor analysis was applied to combine the aforementioned four article measures, article length, number of links, number of sections, and number of images. Parallel analysis suggested that the number of components was 1, with length being the most representative measure (see Figure 1).

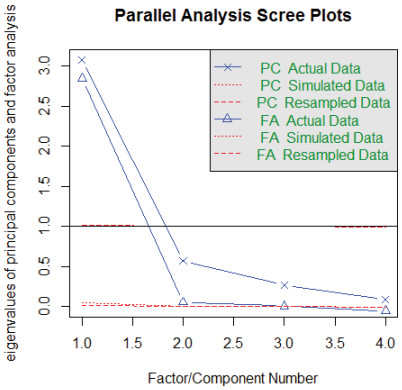


Figure 1: Plot for the factor analysis of the article metric

2.3 Textual analysis

Lead sections of Wikipedia articles were analyzed via two lexicon-based analyzers.

2.3.1 The Hu and Liu approach. This method measures the number of positive and negative words by comparing each word in the text to a list of English positive and negative words [8]. We used this approach in Step 1 to identify the number of positive and negative words in the lead sections.

2.3.2 LIWC2015. LIWC2015 is a software application that provides percentages for every category by comparing each word in the given text with its own dictionary [12]. LIWC2015 was utilized in Step 1 to identify the percentage of positive and negative words in the lead sections. In Step 2, seven main and

seven sub-categories of LIWC psychological constructs were taken into account.

3 RESULTS

3.1 Step 1 – Positivity and negativity of the lead sections

We first measured and compared the positivity and negativity in the lead sections (RQ1). Number of positive words (M = 2.41, SD = 4.53) and negative words (M = 1.55, SD =3.65) differed from each other, $t(126{,}580) = 37.293$, $p < 0.001$. The percentage of positive words (M = 0.84, SD = 1.41) also outnumbered the percentage of negative words (M = 0.51, SD = 0.98), $t(121{,}960) = 45.54$, $p < 0.001$.

Next, we examined the relationships between article metric and positive and negative content of the lead sections (RQ2). Each content variable positively correlated with the article metric with $p < 0.001$ (see Table 1).

Table 1: Correlations for positive/negative content and the article metric

Categories	*r*
Number of positive words*	0.19
Number of negative words*	0.09
Percentage of positive emotion	0.29
Percentage of negative emotion	0.13

*Note. r: strength of the correlation. *Partial correlation analyses were implemented controlling for the number of total words.*

These results suggest that Wikipedia lead sections contained more positive content. The article metric was related to both the number and percentage of positive and negative words. This is in line with the theories of co-activation that people may experience positive and negative emotional aspects in parallel rather than separately [2]. Positive content was, in general, better related with the article metric. This is supporting evidence for the "Pollyanna hypothesis" that points out a universal human tendency to use more positive words in communication [4].

3.2 Step 2 – Lead sections based on psychological processes

In Step 2, we extended the framework of Wikipedia content and article metric beyond positivity and negativity. Therefore, we implemented more detailed analysis with LIWC categories of psychological processes (RQ3). Seven main categories were significantly correlated with the article metric, correlation coefficients ranging from r = -0.01 to 0.32 with $p < 0.001$ (see Table 2).

In order to get a more parsimonious explanation on the data, we broke down the top two correlated categories, drives and affective processes, to their subcategories (see Table 3). Our model that included seven subcategories accounted for 17% of the variance in the article metric, $R^2 = 0.17$, $F(1, 63281) = 1937$, $p < 0.001$.

Table 2: Results of correlation analyses between LIWC main categories and article metric

Categories	*r*
Drives	0.32
Affective processes	0.25
Social processes	0.18
Perceptual processes	0.14
Cognitive processes	0.14
Biological processes	0.11
Relativity	-0.01

Note. r: strength of the correlation.

With Step 2, we managed to demonstrate that psychological content existed in the lead sections such that all LIWC psychological categories were related to the article metric. Among those, sub-categories of drives (achievement, reward, power, affiliation, and risk) and affective processes (positive and negative emotion) were found to have significantly predicted the article metric. Precisely, when the lead sections included these content types, the articles got longer and contained more links, sections, and images.

4 CONCLUSION

This study provides an insight on the association between sentimental and psychological orientation of the Wikipedia lead sections and the articles' basic measures in the Wikipedia community of the society portal. We found that this community is interested in both positive and negative content, though, by putting slightly more attention on positive content. The fact that drive and affective states were the most prominent psychological content shows that the community produces more article features related to challenging notions (achievements, rewards, power) that mainly direct people's life outcomes [15] and emotional experiences [14].

Although the results are limited in terms of revealing small statistical values, our study still provides empirical support for the understanding of how the Wikipedia communities create articles in relation to lead sections' content characteristics. Basing upon these findings, our next plan is to extend the investigation to the entire society portal of the English Wikipedia to get a full grasp on the community's interests. We also plan to include different versions of the articles (e.g., from two different time points) to see whether we could draw inferences predicting article features based on content characteristics. Conducting such studies by adopting a psychological perspective could significantly contribute to Wikipedia research given its vital role in individual and collective knowledge processes [5].

Table 3: Correlations between the top correlated LIWC sub-categories and the article metric

Categories	*r*	LIWC examples	Article examples
Drives			
Achievement	0.308	win, success	Academy Award for Best Visual Effects
Reward	0.295	prize, benefit	Pulitzer Prize for Breaking News Photography
Power	0.276	superior, bully	Child Soldiers International
Affiliation	0.228	ally, friend	Social Democratic Party
Risk	0.134	danger, doubt	Economy of Suriname
Affective processes			
Positive emotion	0.290	love, nice	Academy Award for Best Story
Negative emotion	0.128	hurt, ugly	Battle of Cape St Vincent (1797)

Note. r: strength of the correlation. Article examples include a particularly high amount of words in the respective category.

ACKNOWLEDGMENTS

This work is funded by the project *Analytics for Everyday Learning (AFEL)* within the EU Horizon 2020 program.

REFERENCES

[1] Berger, Jonah, and Kl Milkman. 2013. "Emotion and Virality: What Makes Online Content Go Viral?" *GfK Marketing Intelligence Review* 5 (1):14–23. http://www.gfkmir.com/issues/previous-issues/vol-5-no-1-2013/art3-no1-2013.html.

[2] Berrios, Raul, Peter Totterdell, and Stephen Kellett. 2015. "Eliciting Mixed Emotions: A Meta-Analysis Comparing Models, Types and Measures." *Frontiers in Psychology* 6 : 428. https://doi.org/10.3389/fpsyg.2015.00428.

[3] Blumenstock, Joshua E. "Size matters: word count as a measure of quality on wikipedia." In Proceedings of the 17th international conference on World Wide Web, pp. 1095-1096. ACM, 2008.

[4] Boucher, Jerry, and Charles E. Osgood. 1969. "The Pollyanna Hypothesis." *Journal of Verbal Learning and Verbal Behavior* 8 (1):1–8. https://doi.org/10.1016/S0022-5371(69)80002-2.

[5] Cress, Ulrike, and Joachim Kimmerle. 2008. "A Systemic and Cognitive View on Collaborative Knowledge Building with Wikis." *International Journal of Computer-Supported Collaborative Learning* 3 (2):105–22. https://doi.org/10.1007/s11412-007-9035-z.

[6] Ferron, Michela, and Paolo Massa. 2012. "Psychological Processes Underlying Wikipedia Representations of Natural and Manmade Disasters." *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration - WikiSym 2012*, 1–10. https://doi.org/10.1145/2462932.2462935.

[7] Greving, Hannah, Aileen Oeberst, Joachim Kimmerle, and Ulrike Cress. 2017. "Emotional Content in Wikipedia Articles on Negative Man-Made and Nature-Made Events." *Journal of Language and Social Psychology* 0 (0):0261927X17717568. https://doi.org/10.1177/0261927X17717568.

[8] Hu, Minqing, and Bing Liu. 2004. "Mining and Summarizing Customer Reviews." *Proceedings of the 2004 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '04*, 168. https://doi.org/10.1145/1014052.1014073.

[9] Iosub, Daniela, David Laniado, Carlos Castillo, Mayo Fuster Morell, and Andreas Kaltenbrunner. 2014. "Emotions under Discussion: Gender, Status and Communication in Online Collaboration." *PLoS ONE* 9 (8). https://doi.org/10.1371/journal.pone.0104880.

[10] Liu, Bing. 2012. "Sentiment Analysis and Opinion Mining," no. May:1–108. https://doi.org/10.2200/S00416ED1V01Y201204HLT016.

[11] Oeberst, Aileen, Ina von der Beck, Mitja D. Back, Ulrike Cress, and Steffen Nestler. 2017. "Biases in the Production and Reception of Collective Knowledge: The Case of Hindsight Bias in Wikipedia." *Psychological Research*. Springer Berlin Heidelberg, 1–17. https://doi.org/10.1007/s00426-017-0865-7.

[12] Pennebaker, James W, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. "*The Development and Psychometric Properties of LIWC2015*." Austin, TX: University of Texas at Austin.

[13] Pennebaker, James W, Matthias R Mehl, and Kate G Niederhoffer. 2003. "Psychological Aspects of Natural Language Use: Our Words, Our Selves." *Annual Review of Psychology* 54 (1):547–77. http://arxiv.org/abs/1611.08945.

[14] Rafaeli, Sheizaf, and Yaron Ariel. 2008. "Online Motivational Factors: Incentives for Participation and Contribution in Wikipedia." *Psychological Aspects of Cyberspace: Theory, Research, Applications*, 243–67. https://doi.org/10.1017/CBO9780511813740.012.

[15] Siegling, Alex B., and K. V. Petrides. 2016. "Drive: Theory and Construct Validation." *PLoS ONE* 11 (7):1–27. https://doi.org/10.1371/journal.pone.0157295.

[16] Sundin, Olof. 2011. "Janitors of Knowledge: Constructing Knowledge in the Everyday Life of Wikipedia Editors." *Journal of Documentation* 67 (5):840–62. https://doi.org/10.1108/00220411111164709.

[17] Tausczik, Yla R., and James W. Pennebaker. 2010. "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods." *Journal of Language and Social Psychology* 29 (1):24–54. https://doi.org/10.1177/0261927X09351676.

[18] Wagner, Claudia, Eduardo Graells-Garrido, David Garcia, and Filippo Menczer. 2016. "Women through the Glass Ceiling: Gender Asymmetries in Wikipedia." *EPJ Data Science* 5 (1). Wagner et al. https://doi.org/10.1140/epjds/s13688-016-0066-4.

[19] Wilkinson, D. M., and Huberman, B. A. (2007). Assessing the value of coooperation in Wikipedia, 1–14. https://doi.org/citeulike-article-id:1129224

[20] Wilson, Adam M, and Gene E Likens. 2015. "Content Volatility of Scientific Topics in Wikipedia: A Cautionary Tale." *PLoS ONE* 10 (8):10–14. https://doi.org/10.1371/journal.pone.0134454.

[21] Yenikent, Seren, Peter Holtz, and Joachim Kimmerle. 2017. "The Impact of Topic Characteristics and Threat on Willingness to Engage with Wikipedia Articles: Insights from Laboratory Experiments." *Frontiers in Psychology* 8 (1960):1–11. https://doi.org/10.3389/fpsyg.2017.01960.

[22] Zesch, Torsten, Christof Müller, and Iryna Gurevych. Extracting Lexical Semantic Knowledge from Wikipedia and Wiktionary." In *LREC*, vol. 8, no. 2008, pp. 1646-1652. 2008.

Exploiting the Web of Data for the creation of mobile apps by non-expert programmers

Tatiana Person
Department of Computer Engineering, University of Cádiz
Puerto Real (Cádiz), Spain
tatiana.person@uca.es

Iván Ruiz-Rube
Department of Computer Engineering, University of Cádiz
Puerto Real (Cádiz), Spain
ivan.ruiz@uca.es

Juan Manuel Dodero
Department of Computer Engineering, University of Cádiz
Puerto Real (Cádiz), Spain
juanma.dodero@uca.es

ABSTRACT

The incorporation of specific mobile applications in various disciplines can be very useful. However, the content of these applications may be subject to updates that the developer should perform manually. The use of linked data sources can be a possible solution to this problem, making the content of the applications dynamically updated. Nevertheless, the development of mobile applications capable of including these features is not trivial for a user who does not have adequate programming skills. In this paper, an extension for a mobile application authoring environment based on MIT App Inventor 2 is proposed. This tool provides users with a set of tools to query linked data sources. Finally, this tool will be evaluated by developing a mobile application that will query the database exposed by Wikidata. The aim of the application is to help with the learning of country flags.

KEYWORDS

mobile apps, authoring tools, wikidata, open data, RDF, blockly

1 INTRODUCTION

Nowadays, according to Ditendria's 2017 report [1], 66% of the world population has a smartphone. In addition, the use of mobile applications means 60% of the time spent in the digital world. Mobile applications are, hence, becoming increasingly important in people's daily lives. In existing digital content repositories, such as Google Play Store or App Store, we can find mobile applications of many different topics: communication between people, entertainment, lifestyle control or device monitoring. Mobile applications have also emerged for different educational purposes: explanation of specific topics or concepts, student evaluation, laboratory experiments, collaborative resolution of exercises, language learning, etc.

In several works, [2], [3] and [4], the authors suggest the interest in expanding the availability of content in linked data format for analytical web science and so that it is available for external applications to use. On the other hand, the development of mobile applications that publish specific content depends on the availability and updating of this content. The use of linked data sources can be useful to automatically provide interesting and updated content in these applications. It also reduces the effort invested by the developer to keep content up-to-date and avoiding users to update the application on their mobile devices [5]. However, it is important to note the difficulty of querying the content of linked data sources for users who do not have in-depth programming skills. In this sense, is essential to use environments that allows these users to make use of this type of technologies.

This work presents a set of extensions included in a mobile application development environment based on MIT App Inventor 2. This open-source platform, developed by Google and MIT, allows new users in programming to build applications for Android devices. The goal of this extensions is to facilitate the query of linked data sources from this environment and the generated mobile applications.

2 CONTEXT

2.1 Mobile application authoring tools

Authoring tools are computer applications that facilitate the creation, publication and management of multiple materials in digital format. These tools allow non-expert users to automate the software development processes. For example, with Google Forms, users who are not web developers can easily create web surveys to share and later collect data.

There are several authoring tools that use a visual language for creating applications, such as: Scratch[1], MIT App Inventor[2], Pocket Code[3] or VEDILS[4]. On the other hand, there are other tools that use a textual language, such as: Microsoft Touch Develop[5], Upplication[6], GameSalad[7] or Alice[8].

The analyses conducted in [6] indicate that people without programming knowledge evaluate the structures provided by textual programming languages as unintuitive. Similarly, analyses conducted in [7] claims that people who have started programming using Scratch have a smaller learning curve when faced with programming languages with textual syntax such as C# or Java. Finally, it is important to note that block-based languages have been used successfully in multiple initiatives for the introduction of school-based programming, such as *One Hour of Code*[9]. The use of authoring tools that propose a visual syntax is more indicated for people without technical skills.

[1] https://scratch.mit.edu
[2] http://appinventor.mit.edu
[3] https://share.catrob.at/pocketcode/
[4] http://vedils.uca.es
[5] https://www.touchdevelop.com
[6] https://www.upplication.com
[7] http://gamesalad.com
[8] https://www.alice.org
[9] https://code.org/learn

2.2 Linked Open Data

The concept of Open Data can be defined as data that "anyone can freely access, use, modify, and share for any purpose"[10]. Furthermore, the combination of Web content with semantic descriptions creates an interconnected information structure known as Linked Data [8]. In this type of content, Universal Resource Identifiers (URIs) should be used not only to identify resources as web pages, but also real objects and abstract concepts. These URIs also allow us to retrieve information related to the referenced resource. On the other hand, a single model should be used for the description of resources: the Resource Description Framework (RDF) standard. In addition, the links between resources (RDF triples) must be defined [9]. Finally, this type of content allows its consultation using the SPARQL Protocol and RDF Query Language (SPARQL).

However, these formats and languages to query these contents are not easy to learn for users who are not experts in programming.

2.3 Support of RDF queries from authoring tools

There are some works in the literature related to how to perform RDF queries from block-based environments. First, an extension built on top of MIT App Inventor 2 to issue linked data queries is presented in [10]. These RDF queries are configured in text format from the block editor. This feature allows expert users to define ad-hoc queries, however, users without technical skills may find it more difficult. Another related work is presented in [11]. This work has been also implemented on the basis of MIT App Inventor 2. In this case, the authors propose a feature to auto-complete terms to deal with ontologies with a large number of terms. Finally, in [12] the authors present a tool based on Blockly[11] to make RDF queries using block editor at low level. Nevertheless, its usage by people without technical skills seems still challenging.

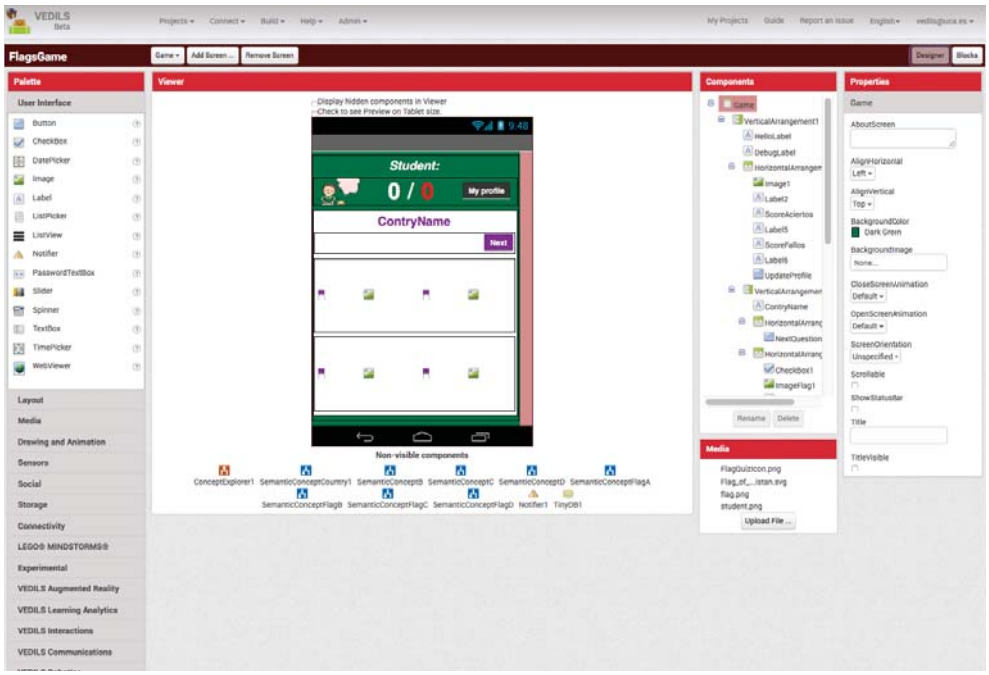
3 RDF QUERIES FROM BLOCK-BASED ENVIRONMENTS

VEDILS is an environment based on MIT App Inventor 2 to easily develop multimodal and interactive learning scenarios. The platform includes a view where the user can design the user interface (see Figure 1a) and a view (Blockly-based editor) where they can define the behavior of the elements included in the applications (see Figure 1b). VEDILS provides a set of additional features that can be integrated with those already provided by MIT App Inventor 2. Features such as augmented reality, virtual reality, gestural interaction and learning analytics, among others are available. Two specific components have been developed to support this research: *ConceptExplorer* and *SemanticConcept*. The purpose of the *ConceptExplorer* component is to retrieve all the existing concepts in an ontology, whereas *SemanticConcept* is intended to load the properties of a given resource. The most relevant features of the above components are described below:

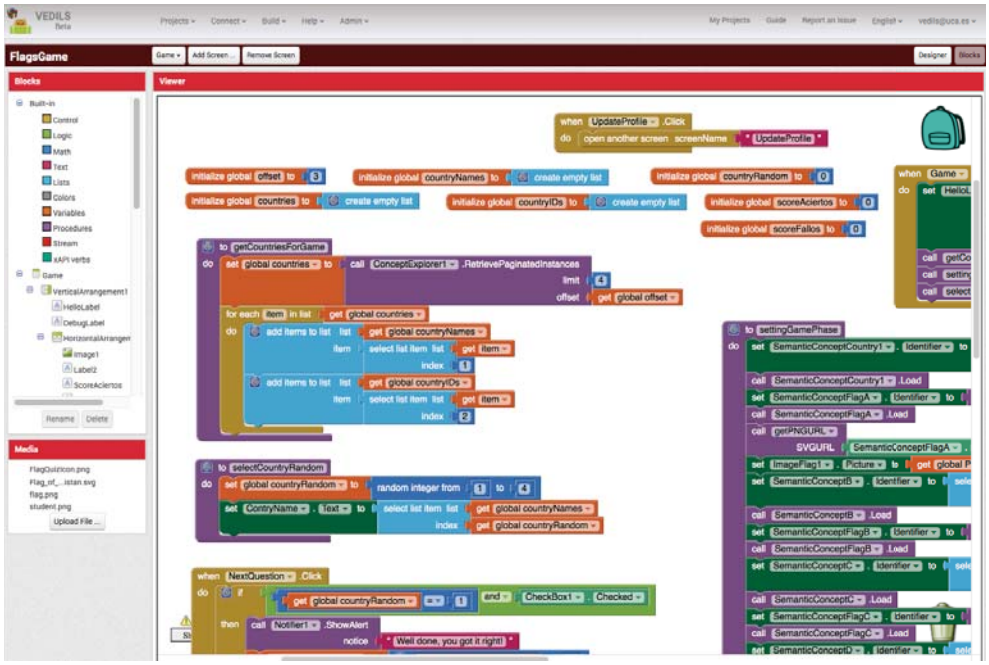
- **ConceptExplorer features** (see Figure 2):
  - *RetrieveAncestors*: Returns the concepts that are ancestors of the selected concept. This function allows you to search for the hierarchical structure of the ontology. The RDF triple used to perform the query is: *?p <conceptURI> rdfs:subClassOf ?p.*
  - *RetrieveDescendants*: Like the above function, it returns the concepts that are descendants of the selected concept. The RDF triple used to perform the query is: *?p rdfs:subClassOf <conceptURI>.*
  - *RetrieveInstances*: Returns the list of instances belonging to the selected concept. The RDF triple used to perform the query is: *?p rdf:type <conceptURI>.*
  - *RetrieveProperties*: Returns the list of properties belonging to the selected concept. The RDF triple used to perform the query is: *?s a <conceptURI> ; ?p ?o.*
- **SemanticConcept features** (see Figure 3):
  - *Identifier*: Specifies the URI which identifies the instance of the concept in the ontology.
  - *AvailableProperties*: Obtains the list of existing properties in the selected instance of the concept. With this function, the developer can know the name of the existing properties in order to get their values. The RDF triple used to perform the query is: *?s a <conceptURI> ; ?p ?o.*
  - *Load*: Downloads the values of all the properties of the concept selected with its identifier from the RDF endpoint. The RDF triple used to perform the query is: *<identifier> ?p ?v UNION ?v ?p <identifier>.*
  - *RetrieveProperty*: Returns the value of a given property belonging to the loaded instance of the concept. The property will be selected through a drop-down menu that lists all existing properties. This function has multiple variants depending on the data type of the selected property. The selectable data types are: URIs, list of URIs, numbers, strings or list of strings.
  - *RetrieveLinkedConcept*: Returns the value of the selected property as a linked concept, that is, it returns its URI. The property will be selected through the dropdowns that show all the existing properties.
- **Common features**:
  - *EndpointRDF*: Specifies the URL of the SPARQL endpoint required for the components.
  - *Classifier*: Specifies the value of the classifier (or type) that must have the selected concept of the ontology.
  - *PreferredLanguage*: Specifies the first preferred language for returning labels of the existing concepts in the ontology.
  - *SecondLanguage*: Specifies the seconds preferred language for returning the labels of the existing concepts in the

[10] http://opendefinition.org
[11] https://developers.google.com/blockly/

ontology. The value selected in this property is only used if the first language is not available in the selected ontology.



(a) Design view of VEDILS



(b) Blockly-based editor of VEDILS

Figure 1: Web interfaces of VEDILS for the creation of mobile applications

4 CASE STUDY

Wikidata is a free and open knowledge base that can be read and edited by both humans and machines. Wikidata acts as central storage for the structured data of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wikisource, and others [12]. The development of applications that issue queries to the Wikidata contents can be interesting because this content is updated and created by a large community.

In order to evaluate our proposal for facilitating the development of linked data apps, we have developed a mobile application using VEDILS that implements a game to boost the learning of country flags (see Figure 4). Because the information used in the application is obtained from Wikidata, we do not have to manually edit the list of countries nor download images of their flags. As a result, the development of this application was relatively straightforward. The developed application is available in Google Play Store[13] for download.

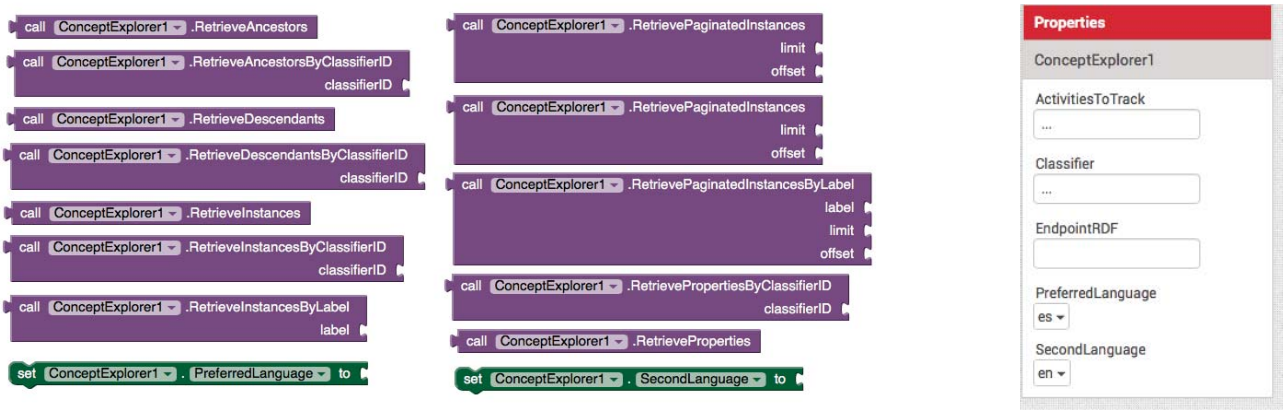
4.1 Application features

The mobile application provides the following functions:

- *Update the user profile*: From the "My profile" button, the user can enter your name and e-mail address. Once registered, it will be used in all sessions of the application.

[12] https://www.wikidata.org/wiki/Wikidata:Main_Page

[13] https://play.google.com/store/apps/details?id=appinventor.ai_vedils.FlagsGame



(a) Blocks

(b) Properties

Figure 2: Blocks and properties of *ConceptExplorer* component

(a) Blocks

(b) Properties

Figure 3: Blocks and properties of *SemanticConcept* component

- *Select the flag of the country*: The application will show the name of a country and we must select the corresponding flag. The selection will be made using the checkboxes shown next to each flag.
- *View the obtained score*: During the course of the game we can view the score obtained. Each time we correctly choose a flag, the application will add one point to our total score.

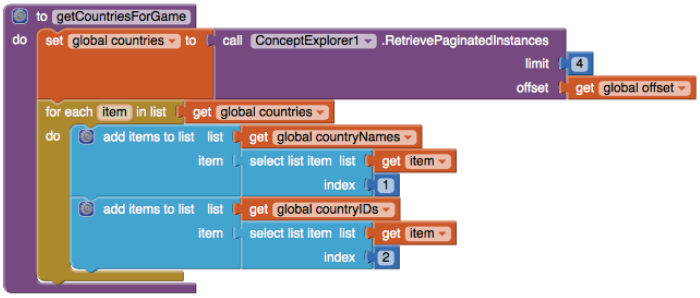
Figure 4: Screenshot of the mobile application developed to guess the flags of the countries.

4.2 Application design

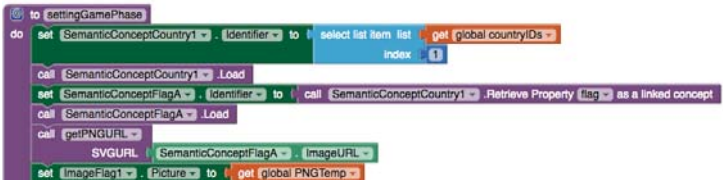
In order to develop the application, the following steps were defined in the app:

- *Configure Wikidata endpoint*: First, the value of the property "EndpointRDF" with the Wikidata SPARQL endpoint was defined. Subsequently, we selected the identifier of the concept "country"[14] provided by Wikidata.
- *Select countries for each phase of the game* (see Figure 5a). To implement the functionality, the instances of the concept "country" must be selected in sets of four elements. To do this, we have used the "RetrievePaginatedInstances" function of the *ConceptExplorer* component with a value of 4 for the limit parameter.
- *Select the flag concept related to the country concept* (see Figure 5b). To show the flag of each country we have used the "flag"[15] linked concept associated to the Wikidata's country concept. With this property we obtain the URL of each country's flag image and then it is render using the *Image* component (included in the components of MIT App Inventor 2).

[14]https://www.wikidata.org/wiki/Q6256
[15]https://www.wikidata.org/wiki/Property:P163



(a) Blocks to retrieve from Wikidata the countries in sets of 4



(b) Blocks to load the flag of one of the countries

Figure 5: Blocks used for configuring the functionality of the mobile application to learn country flags

5 CONCLUSION

This paper presents a set of tools for the querying of linked data sources in a easy way. These tools have been included in a visual authoring environment of mobile applications, VEDILS, which is based on MIT App Inventor 2. The main objective of these tools is to enable novice users in programming to create mobile applications enhanced with data from linked data sources. The advantage of developing applications using this type of information is that mobile apps could automatically have the most up-to-date information reducing effort of the programmer.

A limitation of the presented extensions is that they only allow developers to perform read-only operations on linked data sources. As future work, we will extend the components presented in this work to support the SPARQL Update specification. As a result, in addition to access to the information of different linked data sources, it will be possible to update and insert data into the remote linked data sources from the mobile applications themselves. Furthermore, auto-completion of terms, as proposed in one of the related works, will be also included. This feature will make it easier for non-expert users to find the terms they need in a simpler way. Finally, the tools proposed will be evaluated with end users in order to assess their usefulness and improve them.

ACKNOWLEDGEMENTS

This work has been developed in the VISAIGLE project, funded by the Spanish Ministry of Economy, Industry and Competitiveness with ref. TIN2017-85797-R.

REFERENCES

[1] Ditendria. Informe ditrendia 2017: Mobile en españa y en el mundo, 2017.

[2] Juan Manuel Dodero, Iván Ruiz-Rube, Manuel Palomo-Duarte, and Juan Vázquez-Murga. Open linked data model revelation and access for analytical web science. In *Research Conference on Metadata and Semantic Research*, pages 105–116. Springer, 2011.

[3] Iván Ruiz-Rube, Carlos M Cornejo, Juan Manuel Dodero, and Vicente M García. Development issues on linked data weblog enrichment. In *Research Conference on Metadata and Semantic Research*, pages 235–246. Springer, 2010.

[4] Danillo R Celino, Luana Vetter Reis, Beatriz Franco Martins, and Vítor E Silva Souza. A framework-based approach for the integration of web-based information systems on the semantic web. In *Proceedings of the 22nd Brazilian Symposium on Multimedia and the Web*, pages 231–238. ACM, 2016.

[5] Anthony I Wasserman. Software engineering issues for mobile application development. In *Proceedings of the FSE/SDP workshop on Future of software engineering research*, pages 397–400. ACM, 2010.

[6] Andreas Stefik and Susanna Siebert. An empirical investigation into programming language syntax. *ACM Transactions on Computing Education (TOCE)*, 13(4):19, 2013.

[7] Michal Armoni, Orni Meerbaum-Salant, and Mordechai Ben-Ari. From scratch to "real" programming. *ACM Transactions on Computing Education (TOCE)*, 14(4):25, 2015.

[8] Christian Bizer, Tom Heath, Kingsley Idehen, and Tim Berners-Lee. Linked data on the web (ldow2008). In *Proceedings of the 17th international conference on World Wide Web*, pages 1265–1266. ACM, 2008.

[9] Ivan Salvadori, Alexis Huf, Ronaldo dos Santos Mello, and Frank Siqueira. Publishing linked data through semantic microservices composition. In *Proceedings of the 18th International Conference on Information Integration and Web-based Applications and Services*, pages 443–452. ACM, 2016.

[10] Oshani Seneviratne, Evan W Patton, Daniela Miao, Fuming Shih, Weihua Li, Lalana Kagal, and Carlos Castillo. Developing mobile linked data applications. In *International Semantic Web Conference (Posters & Demos)*, pages 169–172, 2014.

[11] Fuming Shih, Oshani Seneviratne, Ilaria Liccardi, Evan Patton, Patrick Meier, and Carlos Castillo. Democratizing mobile app development for disaster management. In *Joint Proceedings of the Workshop on AI Problems and Approaches for Intelligent Environments and Workshop on Semantic Cities*, pages 39–42. ACM, 2013.

[12] Paolo Bottoni and Miguel Ceriani. Sparql playground: A block programming tool to experiment with sparql. In *VOILA@ ISWC*, page 103, 2015.

Adapting an open source social bookmarking system to observe critical information behaviour

Simone Kopeinik
Institute of Interactive Systems and Data Science, Graz University of Technology
Graz, Austria
simone.kopeinik@tugraz.at

Almonzer Eskandar
School of Digital Technologies, Tallinn University
Tallinn, Estonia
eskandar.almonzer@gmail.com

Tobias Ley
School of Educational Science, Tallinn University
Tallinn, Estonia
tley@tlu.ee

Dietrich Albert
Institute of Interactive Systems and Data Science, Graz University of Technology
Graz, Austria
dietrich.albert@tugraz.at

Paul Seitlinger
School of Educational Science, Tallinn University
Tallinn, Estonia
pseiti@tlu.ee

ABSTRACT

Constructively dealing with societal problems requires a process of opinion formation that is preceded by a competent and diverse search of information. Alarming are thus, communicative processes that can increasingly be observed in social media. In particular, a tendency of drawing virtual circles around like-minded people seems to characterize users' information behaviour and opinion formation dynamics, leading to separated viewpoints of communicative milieus, often referred to as echo chambers. This development raises concern about the public as a hub of diverse perspectives and also starts entering the agenda setting of educational innovation, which aims to prepare students for a more responsible information behaviour. In order to support such innovation, the goal of the present work is to complement prior research on echo chamber phenomena that has mainly made use of questionnaires to get to know the involved socio-cognitive variables and dynamics. By the example of two application scenarios, we showcase how an adapted social bookmarking system can be applied for a more direct observation technique that derives behavioural indices for variables of interest from log-file recordings. We believe that the observation of students' information seeking behaviour will also inspire the design of teaching strategies, e-learning and learning analytics tools.

CCS CONCEPTS

• **Human-centered computing** → **Web-based interaction**; **Social networking sites**; *User studies*; *Open source software*;

KEYWORDS

critical search, depolarisation, user studies, learning analytics, SemanticScuttle

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 123-4567-24-567/08/06…$15.00
https://doi.org/10.475/123_4

ACM Reference Format:
Simone Kopeinik, Almonzer Eskandar, Tobias Ley, Dietrich Albert, and Paul Seitlinger. 2018. Adapting an open source social bookmarking system to observe critical information behaviour. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, Article 4, 4 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

Recent years have recorded an increase in using Social Networking Sites (SNSs) [5] and other online forums as a space for online discussion, opinion formation and interaction with others. Irrespective of our geographic location, we can gather online to view, share and discuss information in a virtual exchange of opinions and participate in deliberative democracy [12]. During online discussions, people interact with content shared by others, get influenced by this content, and then, influence others through their own interactions [3]. Particular dynamics between user dispositions (e.g., open- vs. closed-mindedness) and content of interaction (e.g., controversial vs. consensual topics) can create a public sphere, a notion coined by Jürgen Habermas. According to Peter Dahlgren, it can be defined as "a constellation of communicative spaces in society that permit the circulation of information, ideas, debates, ideally in an unfettered manner, and also the formation of political will" [2]. Though SNSs were not meant in the first place to support processes of a public sphere, they are assumed to cause inadvertent exposure to political difference [1] and thereby, to support more deliberate decision-making that draws on alternative information sources [4]. In contrast to this positive view of SNSs as a public sphere, other authors contest this scenario of deliberation (e.g., [11]). Specifically, they argue that participants in online discussions show selective attention toward prior viewpoints, mainly engage with like-minded people and exhibit closed-mindedness about alternatives [10]. This brings about opinion formation dynamics, which move people towards extreme positions or attitudes. One major reason for this polarising process is confirmatory search, i.e., the selective exposure to partisan information (e.g., [9, 13]). While, for sure, the tendency to selectively expose ourselves to the opinion of like-minded people was present in the pre- digital world [8], the communicative means

in social media, such as personalized information services, might amplify our biases. Through such technology-enhanced exposure to consonant views, initial doubts continuously give way to a growing confidence in one's own opinion, leading a person to strengthen an initial position and attitude [13]. A prominent cognitive explanation of such confirmatory information search bias is the psychological phenomena of cognitive dissonance [6], according to which people feel stressed when faced with divergent opinions. Political scientists who take this pessimistic perspective on SNSs assume that the functionalities of social media, such as personalized information filter [6], resonate with the human motive of reducing cognitive dissonance and thus, reinforce people in performing confirmatory search. As a consequence, users of a SNS run the risk of getting locked into a perpetual echo chamber [9], a metaphor for an interpersonal phenomenon where other people's opinions become echoes of one's own and start reinforcing instead of challenging prior beliefs (e.g., [11]). In many cases, such self-reinforcement fuels the phenomenon of group polarisation and political extremism [14]. Messages in the daily press about hateful Facebook postings (e.g. [7]) make us aware that such excessively cohesive group dynamics quite often result in emotionalized and derogative stances to alternative viewpoints. Due to this development, it became of public interest to educate people in digital literacy. As a consequence, a discussion of possibilities and means to teaching technological application started that exceeds the traditional level of teaching how to effectively use tools and programmes.

In this paper we introduce the application of a social bookmarking platform to observe and interpret users' learning behaviour on the Web. Our goal is to contribute to a better understanding of underlying socio-cognitive dynamics that either lead to a deliberate, open-minded or a biased, polarised information behaviour [15]. As a demonstration and in a first step, we make use of the present online scenario to cross-validate results of a study that has found a systematic relationship between people's tendency to perform confirmatory search on the Web and polarisation [13]. As this finding, however, is based on data gathered through questionnaires, the question remains open whether a similar pattern can be found if both variables are observed directly by extracting them from log-file recordings. With this, we also aim to contribute to the development of teaching strategies of digital competences in schools. A suitable processing of collected data may lead to design implications for formative assessment and timely interventions.

2 ENVIRONMENT

The environment that we propose is build upon the open source social bookmarking tool *SemanticScuttle*[1]. It constitutes a collaborative platform to collect and share information online. The functionality of the platform has been tailored to monitor users' information search behaviour. So far, it supports mechanisms to observe users' collection of resources, their assessment of the trustworthiness of information, a user's tendency towards polarisation and their manifestations of confirmatory search. This has been realized with adaptations in the platform's range of functionality, in its user interfaces and database and the deployment of logging services.

First, to avoid unnecessary training periods, the original platform was reduced in functionality. Remaining functions allow for collecting, annotating and reflecting on Web resources as well as for browsing through bookmarks that are shared among the users. The *Annotation Interface* and the *Search Interface* were extended as described in the following paragraphs:

2.1 Annotation Interface

To support users' reflection on their Web resources, the *Annotation Interface* was adapted as illustrated in Figure 1. It was designed to enable the observation of students' ability to assess the credibility of information, their tendency of polarisation during information search and information consumption and their ability to embed new concepts into their knowledge representation. Figure 1 illustrates the interface that takes basic information about the resource in input fields labelled with one. It consists of the URL, a name and freely chosen keywords (tags). Tags assigned by a user can be used to observe particular semantics of the opinion formation process.

Marked with two is a slider that asks for the user's perception of trustworthiness towards the selected resource. The slider ranges from 0 ("not at all trustworthy") to 10 ("very trustworthy"). In combination with the resource's URL, this information can be used to better understand users' ability to evaluate the quality of information and information sources.

In the last block marked with three, a set of topic aspects is presented to the user. These aspects vary with the search topic and therefore, can be configured by the site administrator. A bipolar rating scale is given by two sliders, ranging from -3 ("very negative"), over 0 ("neutral") to 3 ("very positive"). The sliders ask for the author and user stance towards single aspects and allows for inferring confirmatory search behaviour and polarisation.

2.2 Search Interface

The *Search Interface* presented in Figure 2 consists of two parts. Marked with one is the keyword search that is natively provided by the system. Beyond that, the environment was extended to enable the browsing of Web resources according to positive and negative attitudes towards pre-defined topic aspects (labelled with two). To this end, each aspect was displayed as a clickable keyword within the two boxes. That means that if a user clicks on "Cyborgization" within the "Pro Arguments", all bookmarks in the system (added by any group member) with a positively indicated author stance towards "Cyborgization" are listed to the user. This search aid enables the observation of collaborative information search behaviour and further contributes to the assessment of users' tendencies towards confirmatory search.

2.3 Technical Facts

SemanticScuttle is implemented in PHP and Java Script. Persistent data is saved in an SQL database. To implement the adaptation of the platform, this database was extended to embed information about topic aspects, user and author opinions. In addition, an apache-solr [2] based log data server was developed that is exposed through a Web service interface. It is in place to receive all user interaction data. This data can further be analysed for research purposes or

[1] https://sourceforge.net/projects/semanticscuttle/

[2] http://lucene.apache.org/solr/

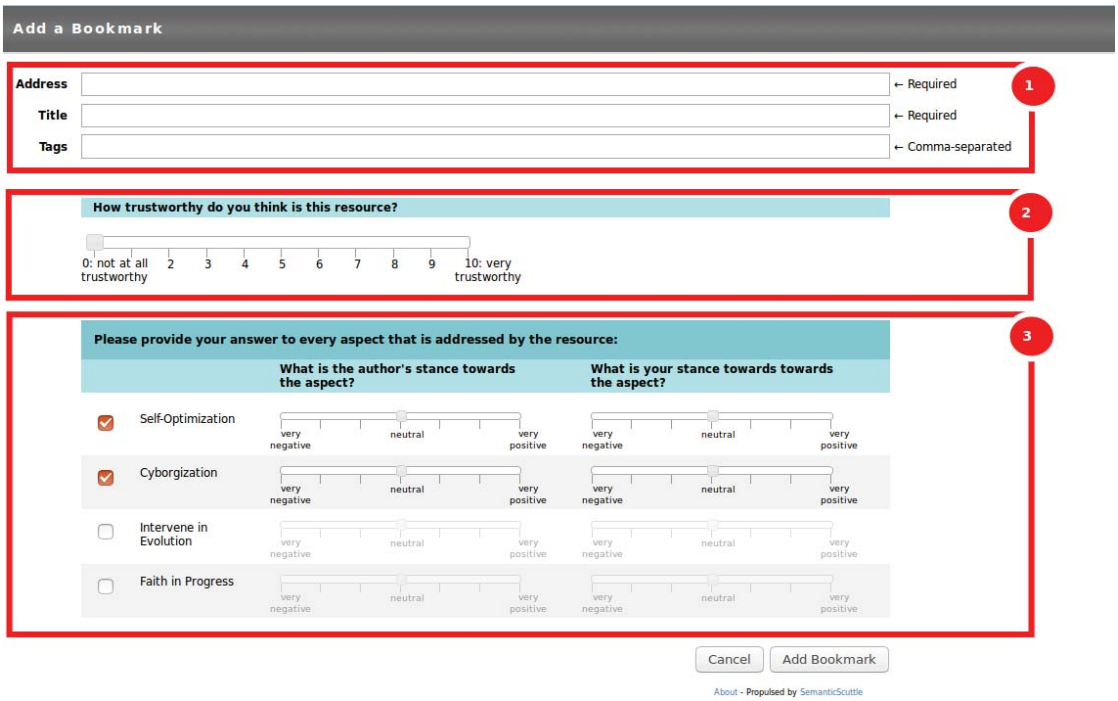


Figure 1: Adapted Annotation Interface.

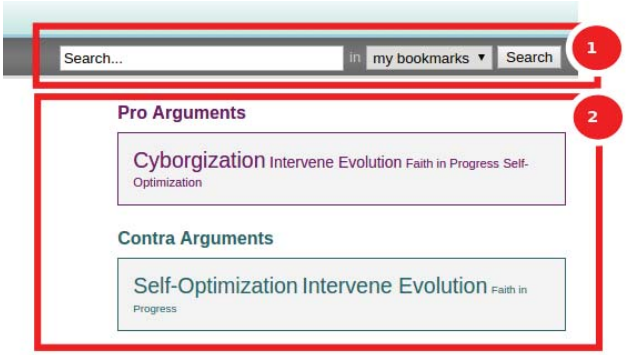


Figure 2: Adapted Search Interface.

accessed continuously to feed into dashboards or other learning analytics tools.

The source code of the adapted *SemanticScuttle* instance and additional background logging services is freely available via github [3].

3 USE CASES

In this section we describe two application scenarios of the software. The first one reports on a pilot approach that was implemented as an online user study with volunteers collecting resources over a period of two weeks. The second example provides insight on a currently running school experiment and its possible application to support formative assessment of students' information search processes.

3.1 Study Procedure

Prior to the study, each participant was provided with a brief description of the study setup and its main research goals. Participants were informed about tasks they had to complete, the data that was gathered and potential privacy concerns. Based on this information, participants or their legal guardians (if applicable) signed a consent form. To ensure data protection and anonymity, users were identified by a pseudonym they created for themselves. For communication purposes we kept a list of email addresses and associated pseudonyms for the duration of the experimental phase.

3.2 Online Pilot Experiment

The first use case presents a pilot study that explores information search behaviour in a controlled online experiment. To this end, 20 participants were recruited through Facebook, of which 13 completed the study.

After signing the consent form, they were instructed to research the topic "transhumanism" online. This topic was selected as we expected it to be controversial while interesting for a wide range of people. At least 20 resources had to be collected continuously over a period of two weeks. Participants added their bookmarks to the provided *SemanticScuttle* instance, using the *Annotation Interface* described in section 2. They had to reflect on the selected resource, assign it to at least one predefined topic facet (i.e., "self-optimization", "cyborgization", "intervene in evolution", "faith in

[3] https://github.com/meins/ReflectiveScuttle

progress") and indicate their personal and the author's stance towards the selected aspects. The resulting workload for a participant was about 20 minutes per day and allowed us to look into different aspects of the opinion formation process, such as confirmatory search or polarisation.

Initial Results. The main research question of this study investigates whether there is a positive correlation between polarisation and confirmatory search. Surprisingly, results showed a significant negative correlation between the two measures indicating that an increase in confirmatory search on a topic is accompanied by a less polarised stance towards the aspect.

This counter-intuitive pattern of results can be explained post-hoc by the fact that most of the participants performing confirmatory search had already started with a rather balanced stance towards different topic aspects. Thus, by collecting resources authored by like-minded people, i.e., by performing a confirmatory search, they had the chance to get to know additional arguments that supported their already developed (balanced) stance and thereby helped further decrease their polarization score (i.e., deviation from a balanced stance).

3.3 Application in the School Context

Currently, the environment is being used in a real-life classroom study. While monitoring students' information search behaviour, the study is part of a participatory design approach to collaboratively develop software that supports the teaching of digital literacy. In this study, a total of 90 high school students between 14 and 18 and three teachers of two schools are taking part. After obtaining parents' informed consent, students attended an initial workshop to become familiar with the problems of echo chambers, filter bubbles and fake news. Also, they were informed about means to evaluate the quality of information. In four school lessons, students researched the topic "global nutrition" under the topic aspects of "genetic engineering", "conservation", "sustainable consumption" and "development aid". The topic and its aspects were selected by the participating teachers in accordance with the curriculum of the age group.

The environment allows us to monitor Web resources students collect and interact with, their ability to evaluate the quality of resources and their tendency towards polarisation and biased search tendencies. In this work, the focus is on understanding students' struggles in online research to allow formative diagnosis and intervention. For instance, if a student assesses a user comment of an SNS as trustworthy, this may indicate a lack in information evaluation skill. Collected log data will further be discussed in teacher workshops to collaboratively design a learning dashboard that will support the formative evaluation of students' information search behaviour.

4 DISCUSSION AND FUTURE WORK

In this paper we have presented an approach to observe information behaviour and opinion formation dynamics directly by using an adapted instance of the open source social bookmarking platform *SemanticScuttle*. In contrast to prior research that explored the relationship between confirmatory search and polarisation (e.g. [15]) on the basis of survey data, the introduced platform offers an environment to investigate this and related phenomena in behavioural observation studies.

Two initial use cases in different contexts (online volunteers, high school students) give an idea of how to apply the system in practice. Our current basis of collected data does not allow for drawing stringent conclusions and the reported results suffer from data loss due to high dropout rates. However, more robust results can be expected from upcoming laboratory studies, where users will operate in a more controlled environment. Also, the presented real-life school study is still in progress, and with 90 participating students promises to improve our understanding of students' information behaviour. Furthermore, the goal of the present and future studies is to derive design implications for further platform development, depolarising discourse services and learning analytics visualizations.

ACKNOWLEDGMENTS

This work is supported by the Austrian Science Fund (FWF): P 27709-G22, TCS-034 and the European Union's Horizon 2020 research and innovation programme under grant agreement No. 669074. We are grateful for the help of Dominik Kowald, Helena Flemming and Marcel Jud in the realization of the experiments.

REFERENCES

[1] Jennifer Brundidge. 2010. Encountering "difference" in the contemporary public sphere: The contribution of the Internet to the heterogeneity of political discussion networks. *Journal of Communication* 60, 4 (2010), 680–700.
[2] Peter Dahlgren. 2005. The Internet, public spheres, and political communication: Dispersion and deliberation. *Political communication* 22, 2 (2005), 147–162.
[3] Cristian Danescu-Niculescu-Mizil, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. No country for old members: User lifecycle and linguistic change in online communities. In *Proceedings of the 22nd international conference on World Wide Web*. ACM, 307–318.
[4] Anna De Liddo and Simon Buckingham Shum. 2013. The Evidence Hub: harnessing the collective intelligence of communities to build evidence-based knowledge. (2013).
[5] Nicole B Ellison et al. 2007. Social network sites: Definition, history, and scholarship. *Journal of computer-mediated Communication* 13, 1 (2007), 210–230.
[6] Leon Festinger. 1954. A theory of social comparison processes. *Human relations* 7, 2 (1954), 117–140.
[7] Klenk Florian. 2016. Boris wollte mich verbrennen. *Falter* 46/16 (11 2016). https://www.falter.at/archiv/wp/boris-wollte-mich-verbrennen
[8] William Hart, Dolores Albarracín, Alice H Eagly, Inge Brechan, Matthew J Lindberg, and Lisa Merrill. 2009. Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological bulletin* 135, 4 (2009), 555.
[9] Robert Huckfeldt, Jeanette Morehouse Mendez, and Tracy Osborn. 2004. Disagreement, ambivalence, and engagement: The political consequences of heterogeneous networks. *Political Psychology* 25, 1 (2004), 65–95.
[10] Michael MacKuen, Jennifer Wolak, Luke Keele, and George E Marcus. 2010. Civic engagements: Resolute partisanship or reflective deliberation. *American Journal of Political Science* 54, 2 (2010), 440–458.
[11] Dimitar Nikolov, Diego FM Oliveira, Alessandro Flammini, and Filippo Menczer. 2015. Measuring online social bubbles. *PeerJ Computer Science* 1 (2015), e38.
[12] Bryan C Semaan, Scott P Robertson, Sara Douglas, and Misa Maruyama. 2014. Social media supporting political deliberation across multiple public spheres: towards depolarization. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 1409–1421.
[13] Natalie Jomini Stroud. 2010. Polarization and partisan selective exposure. *Journal of communication* 60, 3 (2010), 556–576.
[14] Cass R Sunstein. 2001. *Echo chambers: Bush v. Gore, impeachment, and beyond*. Princeton University Press Princeton, NJ.
[15] Yu Wang, Jiebo Luo, Richard Niemi, and Yuncheng Li. 2016. To Follow or Not to Follow: Analyzing the Growth Patterns of the Trumpists on Twitter.. In *News@ ICWSM*.

Current Challenges for Studying Search as Learning Processes

Anett Hoppe
Leibniz Information Centre for Science and Technology (TIB)
Hannover, Germany
anett.hoppe@tib.eu

Peter Holtz
Leibniz Institute of Knowledge Media (IWM)
Tübingen, Germany
p.holtz@iwm-tuebingen.de

Yvonne Kammerer
Leibniz Institute of Knowledge Media (IWM)
Tübingen, Germany
y.kammerer@iwm-tuebingen.de

Ran Yu
L3S Research Centre, Leibniz University Hannover
Hannover, Germany
yu@l3s.de

Stefan Dietze
L3S Research Centre, Leibniz University Hannover
Hannover, Germany
dietze@l3s.de

Ralph Ewerth
Leibniz Information Centre for Science and Technology (TIB)
Hannover, Germany
ralph.ewerth@tib.eu

ABSTRACT

Search of resources and information is among the most frequent activities on the Web. While established information retrieval approaches address the relevance of search results to an information need, the actual learning scope of a user is normally disregarded. Recent research in the *search as learning* (SAL) area has recognized the importance of learning scopes and focused on observing and detecting learning needs.

The article at hand takes a critical look at existing works in SAL and related research disciplines. It aims to give a concise, interdisplinary overview which allows for the deduction of possible directions and necessary actions for prospective research works. It becomes apparent that past research employs a strong emphasis on *textual* resources, neglecting the diversity of online multimedia contents for learning and the impact of multimodal features on the learning process. We argue that exploring multimodal learning resources should be one focus of future SAL projects.

CCS CONCEPTS

• **Information systems** → **Multimedia and multimodal retrieval**; *Web searching and information discovery*; • **Applied computing** → Interactive learning environments;

KEYWORDS

search as learning, learning analytics, multimodal data, multimedia retrieval, educational psychology

ACM Reference format:
Anett Hoppe, Peter Holtz, Yvonne Kammerer, Ran Yu, Stefan Dietze, and Ralph Ewerth. 2018. Current Challenges for Studying Search as Learning Processes. In *Proceedings of Learning and Education with Web Data, Amsterdam, Netherlands, May 2018 (LILE2018),* 4 pages.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

LILE2018, May 2018, Amsterdam, Netherlands

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Information Retrieval research has, for a long time, centered around the concept of an "information need" – a desire to amend a certain identifed lack of information. While fact search is certainly one facet of Web search, recent research points to a multitude of other usage behaviours which are to-date insufficiently supported by technology.

The *Search as Learning* domain examines one of these alternative search facets, that is, search sessions that are related to a learning intent. It relies on the assumption that current search systems are regularly used to access and internalise new knowledge, related to a defined (conscious or unconscious) learning objective. This is reflected in the prevalence of *informational* search intents, which, contrary to transactional or navigational search intents, imply a dedicated learning intent [9]. Technologies developed under the SAL paradigm will have to roughly fulfil the following goals: (a) supporting the users in their learning tasks through an enhanced retrieval and ranking process; (b) enable the accurate detection and prediction of learning needs and scopes, e.g. whether a user intends to acquire declarative or procedural knowledge, as well as respective knowledge gains during search relying on available data (e.g. queries, resource features, behavioural and navigational data); (c) addressing and evaluating both general Web search scenarios as well as semi-informal learning settings that involve search for scholarly information, specifically literature and videos in digital library portals (e.g. the TIB's web portals[1]). In this paper, we give a brief overview of research results regarding SAL processes from the fields of information retrieval and educational psychology. Based on the current state of the art, we identify current challenges for SAL research (Section 3). In particular, we identify a lack of consideration of multimodal resources in SAL, even though their usefulness is supported by multimedia learning research. This article limits its scope to research on individual learning, for the sake of conciseness. Anyhow, the integration of the insights provided by research on collaborative and social learning into SAL systems is, indeed, another interesting research topic.

The remainder of the paper is organised as follows. Related work is briefly reviewed in Section 2. In Section 3, we describe the main challenges of future work in the SAL field from our perspective, while some conclusions are drawn in Section 4.

[1] https://av.tib.eu/

2 RELATED WORK

As stated by Ghosh et al. [17], information science research contains a number of studies which seek to connect search processes and knowledge building (e.g. [13, 19]). Anyhow, it is only recently that research efforts from information science, educational psychology, learning analytics and information retrieval are united with the objective of improving learning support in information retrieval systems. The following paragraphs give a short summary of the main contributions which constitute today's understanding of SAL processes. Subsequently, Section 3 deduces some of the main research gaps.

2.1 SAL in Educational Psychology

Learning can be defined as the act of gaining new or modifying or reinforcing existing knowledge [31]. So far, research has mostly focused on the use of the Internet to gain factual knowledge or to learn about complex, conflicting issues of fragile evidence, that is, conceptual knowledge. Procedural knowledge, i.e., how to perform a certain task, has been hardly examined in the SAL context.

Process model: Commonly, the information seeking process is described as a sequence of processing steps (e.g., [8, 16]): (1) Identifying and defining the information need and generating respective search terms; (2) locating information sources, e.g., web pages, by evaluating and selecting links from search engine result pages (SERPs); (3) evaluating the information presented in web pages; (4) processing and extracting content from web pages identified as useful; and (5) comparing, integrating, and synthesizing information from several resources to prepare the final task outcome (in the user's mind or externally). In particular, step 5 involves a learning component when referring to an internal integration of the retrieved information.

Measuring learning outcomes: The achieved learning is measured as the outcome of the above process. Evaluation methods include counting correctly restored concepts in problem-specific essay tasks or knowledge tests with multiple-choice or true-false items (cf. [38]). However, previous research mostly has focused on learning from textual resources and does not specifically address learning from multimedia data, such as graphical representations or videos. Research from the field of multimedia learning indicates, however, that visual material in addition to text might be beneficial to learning outcomes: [32] find that additional visualisations support, in particular, the learning of procedural knowledge. [5] state that using multimedia, animations and hypertext elements can lead to "deep comprehension of the material", but also lead to problems due to split attention. In their study, adapted animations contribute in particular to the learning of dynamic information. The study presented in [10] suggests that the integration of video material improves learning performance, even in learner types who preferred verbal material over visuals.

Learning success factors: Several studies have examined factors that influence learning success in the processing steps mentioned above. Identified factors include prior domain knowledge (e.g., [37]), personal beliefs with respect to knowledge and how it develops (e.g., [26]; [21, 30]), prior training on evaluating Internet information (e.g., [20, 34, 36]), and usage of alternative search interfaces (e.g., [28]). While research on learning with hypertext and hypermedia systems, i.e., closed learning environments, has shown that an optimal navigational path results in better learning outcomes (e.g., [27]), this remains an open question for Web search scenarios.

2.2 SAL in Information Retrieval

Supporting informal learning has been subject to a plethora of research, whereas learning as an implicit part of search and information retrieval has only recently been recognized [1]. Recent efforts include the automatic identification of users' learning needs and intents from query logs, e.g., declarative or procedural knowledge ([14]). Vakkari [31] provides a well-structured survey of features indicating learning needs as well as user's knowledge and knowledge gain throughout the search process. Other works seek to predict the users' prior domain knowledge as one of the factors correlated to positive learning outcomes: Zhang et al. [39] identify distinctive features in the users' search behavior as predictors of domain knowledge; Cole et al. [11] observe behavioral patterns as reliable indicators; Collins-Thompson et al. [12] find that the usage of intrinsically diverse search queries is positively correlated to increased knowledge gain. A recent study [15] investigates the correlation of search behaviour and users knowledge gain and knowledge state in search sessions across a range of topics, finding only weak correlations with session features but medium correlation with the respective search topics.

Initial efforts aim at integrating the insights of the SAL community into information retrieval systems. Building on the features proposed by Eickhoff et al. [14], Weingart and Eickhoff [35] investigate adapted query expansion and re-ranking techniques in order to improve retrieval results with respect to the users' learning needs. Similarly, Syed and Collins-Thompson [29] examine the effect of keyword density on knowledge gain in language learning tasks.

While several studies underline the positive impact of visual elements on learning processes, e.g., in Web navigation [33] or e-Learning [24], the aforementioned works disregard multimodal aspects and features. In the context of Web search, Karanam et al. [22] show that the assignment of highly relevant pictures to text and hyperlinks significantly reduces the users' efforts to accomplish their search goals. It thus seems likely that the quality and efficiency of SAL processes can be notably increased by considering resource modality aspects and, in particular, retrieving non-textual resources based on the users' learning needs. However, visual elements need to be chosen dynamically based on the learning objective. So far, there are only approaches outside the SAL context that aim at accomplishing this enrichment in an automated way, e.g., Agrawal et al. [3, 4] suggest two different methods based on image metadata and aggregated Web search results, respectively, to enrich textbooks for schools with images. Other proposals address the assignment of relevant videos [2, 23].

3 CHALLENGES

Search as learning is an inherently interdisciplinary research area – current research works unite findings from educational psychology, learning analytics and information retrieval to provide enhanced support for learning tasks during Web search. Given the recentness

of SAL research, it is unsurprising that the review of published works reveals a number of open research questions.

Theoretical frameworks for SAL: . Current SAL research employs a diverse set of theoretical frameworks from different research domains – publications refer, for instance, to Bloom's taxonomy of learning objectives and its derivatives [6, 7] and/or Marchionini's exploratory search paradigm [25]. Anyhow, an integrated view on Search as Learning as an independent concept is still missing. While the currently used frameworks cover important facets of SAL processes, the importance of learning on the Web justfies the development of a unified, theoretical model of Search as Learning itself.

Detection and prediction of learning: A crucial challenge is the detection and understanding of learning and knowledge acquisition in heterogeneous and unstructured online interactions. This includes, for instance, the detection of learning-related search missions, the prediction and classification of learning intents, as well as the prediction of user knowledge and knowledge gain [15] troughout a search mission. However, given that such information is not explicitly provided throughout the search session, SAL research and tools have to consider a wide variety of implicit features observable throughout search sessions, for instance, considering the user interactions, behavioural features, session-related information and multimodal characteristics of the resources used as part of the search process.

Data acquisition: Current works in *Search as Learning* reference the usage of navigational logs, click-through data and eye-tracking experiments as their data sources. Only few of the datasets are openly available for other researchers. The availability of standardised, structured datasets about SAL processes (from lab experiments and collected in the wild) will largely enhance the research landscape – by providing a common base for research and evaluation. Furthermore, data acquisition could use not only controlled laboratory experiments, but also semi-formal settings in crowdsourcing platforms (which would also lead to an important extension of the possible range of participants reached by a study). Structured data acquistion will moreover allow the identification of potential new criteria and features, the discovery of formerly unknown correlations and inter-relationships, and the development of formal, standardised methods for comparative evaluation.

Retrieval and ranking beyond textual resources: Whereas the aforementioned challenges address the understanding and classification of learning throughout search missions, detected learning needs and behaviour have to be supported through dedicated retrieval and ranking processes. For instance, ranking of resources should consider the actual knowledge state of a user and his/her learning intent. In particular, research shows that the inclusion of non-textual resources in educational materials can contribute to the learners' comprehension, internalisation and entertainment. However, SAL research so far has not reached a state where the direct reflection of learning-specific features is reflected through the actual retrieval method. In addition, multimodal resource features are so far under-investigated, despite their relevance for particular learning needs. To provide comprehensive support for different types of learners and learning tasks, interactively (and individually) composed learning materials should include text as well as images and video material of different types. This stream of research – covered on the information retrieval side, for instance, in the domain of multimedia retrieval – has not been tackled in an SAL context, yet.

4 CONCLUSIONS

Aforementioned challenges can only be fully tackled throu collaboration of experts from the related research domains: aquisition of reliable ground truth data involves experiments and quasi-experiments best organised by researchers with psychological background and deep insight in study design, using the full toolkit offered by psychological research. On the other hand, development of predictive models requires knowledge in data analysis and artificial intelligence while expertise in scalable data processing is required to obtain, organise, process and publish collected data to make sure it is reusable across disciplines.

Analysis must be an iterative process – experts from learning analytics can analyse the datasets, discover formerly unknown features of the observed learning processes, discover novel correlations and evaluation measures. Results should be directly fed back in the study design process and validated (or revoked) by further experiments.

Finally, multimodal retrieval should be introduced to SAL research as a novel facet, given that learning research strongly suggests that the inclusion of image and video resources may enhance students' learning outcome. For this purpose, media types and multimodal features have to be included as a feature in the retrieval process (depending on enhanced retrieval of multimodal features and adapted ranking procedures). Some of these challenges will be tackled by the research project "SALIENT: Search as Learning: Investigating, ENhancing and PredicTing Learning during Multimodal Web Search", funded by the Leibniz Assocation in Germany from 2018 to 2021. It is a collaborative research project involving partners from information retrieval (L3S), educational psychology (IWM), and multimedia retrieval (TIB).

5 ACKNOWLEDGEMENTS

This work is financially supported by the Leibniz Association, Germany (Leibniz Competition 2017, funding line "Cooperative Excellence", project SALIENT [K68/2017]).

REFERENCES

[1] Maristella Agosti, Norbert Fuhr, Elaine G. Toms, and Pertti Vakkari. 2013. Evaluation Methodologies in Information Retrieval (Dagstuhl Seminar 13441). *Dagstuhl Reports* 3, 10 (2013), 92–126. https://doi.org/10.4230/DagRep.3.10.92

[2] Rakesh Agrawal, Maria Christoforaki, Sreenivas Gollapudi, Anitha Kannan, Krishnaram Kenthapadi, and Adith Swaminathan. 2014. Mining Videos from the Web for Electronic Textbooks. In *Formal Concept Analysis - 12th International Conference, ICFCA 2014, Cluj-Napoca, Romania, June 10-13, 2014. Proceedings (Lecture Notes in Computer Science)*, Cynthia Vera Glodeanu, Mehdi Kaytoue, and Christian Sacarea (Eds.), Vol. 8478. Springer, 219–234. https://doi.org/10.1007/978-3-319-07248-7_16

[3] Rakesh Agrawal, Sreenivas Gollapudi, Anitha Kannan, and Krishnaram Kenthapadi. 2011. Data mining for improving textbooks. *SIGKDD Explorations* 13, 2 (2011), 7–19. https://doi.org/10.1145/2207243.2207246

[4] Rakesh Agrawal, Sreenivas Gollapudi, Anitha Kannan, and Krishnaram Kenthapadi. 2011. Enriching textbooks with images. In *Proceedings of the 20th ACM Conference on Information and Knowledge Management, CIKM 2011, Glasgow, United Kingdom, October 24-28, 2011*, Craig Macdonald, Iadh Ounis, and Ian Ruthven (Eds.). ACM, 1847–1856. https://doi.org/10.1145/2063576.2063843

[5] F. Amadieu, J. Lemarié, and A. Tricot. 2017. How may multimedia and hypertext documents support deep processing for learning? *Psychologie Française* 62, 3 (2017), 209 – 221. https://doi.org/10.1016/j.psfr.2015.04.002 Cognition et multimédia : les atouts du numérique en situation d'apprentissage.

[6] L.W. Anderson, D.R. Krathwohl, and B.S. Bloom. 2001. *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. Longman. https://books.google.de/books?id=EMQlAQAAIAAJ

[7] Benjamin S Bloom et al. 1956. Taxonomy of educational objectives. Vol. 1: Cognitive domain. *New York: McKay* (1956), 20–24.

[8] Saskia Brand-Gruwel, Iwan Wopereis, and Amber Walraven. 2009. A descriptive model of information problem solving while using internet. *Computers & Education* 53, 4 (2009), 1207–1217. https://doi.org/10.1016/j.compedu.2009.06.004

[9] Andrei Broder. 2002. A taxonomy of web search. In *ACM Sigir forum*, Vol. 36. ACM, 3–10.

[10] Chih-Ming Chen and Ying-Chun Sun. 2012. Assessing the effects of different multimedia materials on emotions and learning performance for visual and verbal style learners. *Computers & Education* 59, 4 (2012), 1273–1285. https://doi.org/10.1016/j.compedu.2012.05.006

[11] Michael J. Cole, Jacek Gwizdka, Chang Liu, Nicholas J. Belkin, and Xiangmin Zhang. 2013. Inferring user knowledge level from eye movement patterns. *Information Processing & Management* 49, 5 (2013), 1075 – 1091. https://doi.org/10.1016/j.ipm.2012.08.004

[12] Kevyn Collins-Thompson, Soo Young Rieh, Carl C. Haynes, and Rohail Syed. 2016. Assessing Learning Outcomes in Web Search: A Comparison of Tasks and Query Strategies. In *Proceedings of the 2016 ACM Conference on Human Information Interaction and Retrieval, CHIIR 2016, Carrboro, North Carolina, USA, March 13-17, 2016*, Diane Kelly, Robert Capra, Nicholas J. Belkin, Jaime Teevan, and Pertti Vakkari (Eds.). ACM, 163–172. https://doi.org/10.1145/2854946.2854972

[13] B. Dervin. 1983. *An Overview of Sense-making Research: Concepts, Methods, and Results to Date*. The Author. https://books.google.de/books?id=wInhAAAAMAAJ

[14] Carsten Eickhoff, Jaime Teevan, Ryen White, and Susan T. Dumais. 2014. Lessons from the journey: a query log analysis of within-session learning. In *Seventh ACM International Conference on Web Search and Data Mining, WSDM 2014, New York, NY, USA, February 24-28, 2014*, Ben Carterette, Fernando Diaz, Carlos Castillo, and Donald Metzler (Eds.). ACM, 223–232. https://doi.org/10.1145/2556195.2556217

[15] Ujwal Gadiraju, Ran Yu, Stefan Dietze, and Peter Holtz. 2018. Analyzing Knowledge Gain of Users in Informational Search Sessions on the Web. In *2018 ACM on Conference on Human Information Interaction and Retrieval (CHIIR)*. ACM.

[16] Peter Gerjets, Yvonne Kammerer, and Benita Werner. 2011. Measuring spontaneous and instructed evaluation processes during Web search: Integrating concurrent thinking-aloud protocols and eye-tracking data. *Learning and Instruction* 21, 2 (2011), 220 – 231. https://doi.org/10.1016/j.learninstruc.2010.02.005 Special Section I: Solving information-based problems: Evaluating sources and information Special Section II: Stretching the limits in help-seeking research: Theoretical, methodological, and technological advances.

[17] Souvick Ghosh, Manasa Rath, and Chirag Shah. 2018. Searching as Learning: Exploring Search Behavior and Learning Outcomes in Learning-related Tasks. In *Proceedings of the 2018 Conference on Human Information Interaction&Retrieval, CHIIR 2018, New Brunswick, NJ, USA, March 11-15, 2018*, Chirag Shah, Nicholas J. Belkin, Katriina Byström, Jeff Huang, and Falk Scholer (Eds.). ACM, 22–31. https://doi.org/10.1145/3176349.3176386

[18] Jacek Gwizdka, Preben Hansen, Claudia Hauff, Jiyin He, and Noriko Kando (Eds.). 2016. *Proceedings of the Second International Workshop on Search as Learning, SAL 2016, co-located with the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2016, Pisa, Italy, July 21st, 2016*. CEUR Workshop Proceedings, Vol. 1647. CEUR-WS.org. http://ceur-ws.org/Vol-1647

[19] Peter Ingwersen. 1996. Cognitive Perspectives of Information Retrieval Interaction: Elements of a Cognitive IR Theory. *Journal of Documentation* 52, 1 (1996), 3–50. https://doi.org/10.1108/eb026960

[20] Yvonne Kammerer, Dorena G. Amann, and Peter Gerjets. 2015. When adults without university education search the Internet for health information: The roles of Internet-specific epistemic beliefs and a source evaluation intervention. *Computers in Human Behavior* 48 (2015), 297–309. https://doi.org/10.1016/j.chb.2015.01.045

[21] Yvonne Kammerer, Ivar Bråten, Peter Gerjets, and Helge I. Strømsø. 2013. The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during Web search on a medical issue. *Computers in Human Behavior* 29, 3 (2013), 1193–1203. https://doi.org/10.1016/j.chb.2012.10.012

[22] Saraschandra Karanam, Herre van Oostendorp, and Bipin Indurkhya. 2012. Evaluating CoLiDeS + Pic: the role of relevance of pictures in user navigation behaviour. *Behaviour & IT* 31, 1 (2012), 31–40. https://doi.org/10.1080/0144929X.2011.606335

[23] Marios Kokkodis, Anitha Kannan, and Krishnaram Kenthapadi. 2014. Assigning videos to textbooks at appropriate granularity. In *First (2014) ACM Conference on Learning @ Scale, L@S 2014, Atlanta, GA, USA, March 4-5, 2014*, Mehran Sahami, Armando Fox, Marti A. Hearst, and Michelene T. H. Chi (Eds.). ACM, 199–200. https://doi.org/10.1145/2556325.2567880

[24] Els Kuiper, Monique Volman, and Jan Terwel. 2005. The Web as an Information Resource in K-12 Education: Strategies for Supporting Students in Searching and Processing Information. *Review of Educational Research* 75, 3 (2005), 285–328. http://www.jstor.org/stable/3515984

[25] Gary Marchionini. 2006. Exploratory search: from finding to understanding. *Commun. ACM* 49, 4 (2006), 41–46. https://doi.org/10.1145/1121949.1121979

[26] Lucia Mason, Angela Boldrin, and Nicola Ariasi. 2010. Searching the Web to learn about a controversial topic: are students epistemically active? *Instructional Science* 38, 6 (2010), 607–633. https://doi.org/10.1007/s11251-008-9089-y

[27] Ladislao Salmeron, Jose J. Canas, Walter Kintsch, and Inmaculada Fajardo. 2005. Reading Strategies and Hypertext Comprehension. *Discourse Processes* 40, 3 (2005), 171–191. https://doi.org/10.1207/s15326950dp4003_1 arXiv:http://dx.doi.org/10.1207/s15326950dp4003$_1$

[28] Ladislao Salmerón, Laura Gil, Ivar Bråten, and Helge I. Strømsø. 2010. Comprehension effects of signalling relationships between documents in search engines. *Computers in Human Behavior* 26, 3 (2010), 419–426. https://doi.org/10.1016/j.chb.2009.11.013

[29] Rohail Syed and Kevyn Collins-Thompson. 2016. Optimizing Search Results for Educational Goals: Incorporating Keyword Density as a Retrieval Objective, See [18]. http://ceur-ws.org/Vol-1647/SAL2016_paper_21.pdf

[30] Yi-Wen Tu, Meilun Shih, and Chin-Chung Tsai. 2008. Eighth graders' web searching strategies and outcomes: The role of task types, web experiences and epistemological beliefs. *Computers & Education* 51, 3 (2008), 1142–1153. https://doi.org/10.1016/j.compedu.2007.11.003

[31] Pertti Vakkari. 2016. Searching as learning: A systematization based on literature. *J. Information Science* 42, 1 (2016), 7–18. https://doi.org/10.1177/0165551515615833

[32] Erlijn van Genuchten, Katharina Scheiter, and Anne Schüler. 2012. Examining learning from text and pictures for different task types: Does the multimedia effect differ for conceptual, causal, and procedural tasks? *Computers in Human Behavior* 28, 6 (2012), 2209–2218. https://doi.org/10.1016/j.chb.2012.06.028

[33] Herre van Oostendorp, Saraschandra Karanam, and Bipin Indurkhya. 2012. CoLiDeS+ Pic: a cognitive model of web-navigation based on semantic information from pictures. *Behaviour & IT* 31, 1 (2012), 17–30. https://doi.org/10.1080/0144929X.2011.603358

[34] Amber Walraven, Saskia Brand-Gruwel, and Henny P. A. Boshuizen. 2013. Fostering students' evaluation behaviour while searching the internet. *Instructional Science* 41, 1 (2013), 125–146. https://doi.org/10.1007/s11251-012-9221-x

[35] Nino Weingart and Carsten Eickhoff. 2016. Retrieval Techniques for Contextual Learning, See [18]. http://ceur-ws.org/Vol-1647/SAL2016_paper_16.pdf

[36] Jennifer Wiley, Susan R. Goldman, Arthur C. Graesser, Christopher A. Sanchez, Ivan K. Ash, and Joshua A. Hemmerich. 2009. Source Evaluation, Comprehension, and Learning in Internet Science Inquiry Tasks. *American Educational Research Journal* 46, 4 (2009), 1060–1106. https://doi.org/10.3102/0002831209333183 arXiv:http://dx.doi.org/10.3102/0002831209333183

[37] Teena Willoughby, S. Alexandria Anderson, Eileen Wood, Julie Mueller, and Craig Ross. 2009. Fast searching for information on the Internet to use in a learning context: The impact of domain knowledge. *Computers & Education* 52, 3 (2009), 640–648. https://doi.org/10.1016/j.compedu.2008.11.009

[38] Mathew J. Wilson and Max L. Wilson. 2013. A comparison of techniques for measuring sensemaking and learning within participant-generated summaries. *Journal of the American Society for Information Science and Technology* 64, 2 (2013), 291–306. https://doi.org/10.1002/asi.22758

[39] Xiangmin Zhang, Michael J. Cole, and Nicholas J. Belkin. 2011. Predicting users' domain knowledge from search behaviors. In *Proceeding of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2011, Beijing, China, July 25-29, 2011*, Wei-Ying Ma, Jian-Yun Nie, Ricardo A. Baeza-Yates, Tat-Seng Chua, and W. Bruce Croft (Eds.). ACM, 1225–1226. https://doi.org/10.1145/2009916.2010131

Linked Data Generation for Adaptive Learning Analytics Systems

Sven Lieber
sven.lieber@ugent.be

Ben De Meester
ben.demeester@ugent.be

Anastasia Dimou
anastasia.dimou@ugent.be

Ruben Verborgh
ruben.verborgh@ugent.be

Ghent University – imec – IDLab, Department of Electronics and Information Systems
Technologiepark 19
9052 Zwijnaarde, Belgium

ABSTRACT

According to the Learning Analytics (LA) reference model, LA is used to collect, explore and analyze diverse types and interrelationships of data. Specifications like the Experience API (xAPI) work towards interoperability with respect to interrelationship of diverse learning data. Algorithms for adaptive learning could be improved by incorporation of user-related data, not present in learning activities. Linking these user-related data with learning activity data would fully exploit the potential of interrelationships with data. Conventional solutions, as well as current Linked Data-based solutions focus purely on learning activity data, whereas solutions based on Linked Data could be used to integrate data of different domains. We propose a provenance-aware pipeline to transform xAPI learning activity statements to Linked Data. The integration of learning activities with other user data, provides a more complete set of user data, improving an adaptive learning analytics system. We use the proposed pipeline to build a Linked Learning Record Store based on the Resource Description Framework (RDF). SPARQL queries are used to link data about learning activities, enriched with fine-grained exercise descriptions, with data describing the abilities of users. In this paper, we show how Linked Data can be generated from xAPI statements in a streaming approach, based on existing tools and interfaces. Our solution demonstrates the usage of Linked Data to combine learning activity data with user ability data, to get a more complete set of user data aiming to assist in adaptive learning.

1. INTRODUCTION

Learning Analytics aim to collect, explore and analyze diverse types and interrelationships of data [1]. Specifications, such as the Experience API (xAPI)[1], lead towards analysis of different learning experiences. However, valuable data outside the Learning Analytics context are not considered with conventional solutions. Linked Data offers the functionality to analyze interrelationships between Learning Activity and other data.

Abilities of users adapt over time, while certain abilities might cover different domains, e.g., maths or reading. These user-centric skills offer valuable insights regarding learning exercises and analysis of users'performance. By combining these kind of user data with learning activity data, adaptive learning is enabled. Developing a self-learning analytics system for adaptive learning is the purpose of the project LEarning analytics for AdaPtive Support (LEAPS, 2016-2018)[2]. Within LEAPS, multiple educational applications were fitted with xAPI logging, to provide a more integrated analysis. Analytical models can be re-used across different use-cases, due to the fact that the learning activities as well as users skills are modelled as Linked Data. Initial data capture has been performed on 18 classes of the first grade in Flemish schools, in 16 sessions and produced more than 2 million xAPI statements.

In this paper, we propose a data processing layer based on Linked Data on top of commonly used Learning Record Stores (LRS). To achieve that, we propose to use an Extract-Transform-Load (ETL) process to (i) read learning activity data expressed in xAPI from commonly used Learning Record Stores (LRS) in a streaming approach, (ii) apply use-case specific transformations to Linked Data, and (iii) load the Linked Data to a triple store for further consumption.

The remaining of the paper is organized as follows: First, we cover different Linked Data approaches regarding Learning Analytics in Section 2. Secondly, we present our integration model in Section 3 and our proposed Linked Data architecture in Section 4. We demonstrate interlinking of learning activity data with user-related data regarding abilities in Section 5. And finally we discuss our conclusion in Section 6.

2. RELATED WORK

© 2018 ACM. ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

[1]https://xapi.com/
[2]https://www.imec-int.com/en/what-we-offer/research-portfolio/leaps_2

Utilizing Linked Data principles for Learning Analytics is not a novel approach. Rabelo et al. [9] proposed a big data architecture using the xAPI ontology and Linked Data. However, they were using it for validation regarding xAPI specification conformance only, and they did not fully exploit the possibilities of Linked Data for Learning Analytics.

De Nies et al. [4] performed the first steps toward semantic interoperability of xAPI learning statements with other domains. They identified learning activities as provenance and proposed the tool TinCan2PROV[3] to generate provenance-related Linked Data from xAPI statements. They relied on JSON-LD contexts to map xAPI terms to an xAPI ontology based on the specification. We are re-using their JSON-LD context within our proposed ETL process as one transformation step.

Anseeuw et al. [2] created an xAPI extension to store more context information regarding the physical environment of the learning activity. They also extended the work of De Nies et al. [4] concerning the xAPI ontology and mainly used it for reasoning. They defined additional services, based on SPARQL queries, which query learning activity data regarding the result and the provided context information. However, their work does not make use of data other than present in the learning statement.

De Meester et al. [3] created the Semantic ExerCise Interchange Format (SERIF[4]) to semantically describe learning exercises. The learning activity statements of the LEAPS project are described using the *SERIF* format as xAPI extension. Interoperable fine-grained description of exercises offer a variety of ways to interlink with user-related data. One transformation step of our ETL process makes use of the provided JSON-LD context of De Meester et al. [3] to generate fine-grained Linked Data describing the exercises. We are then able to properly link user-related abilities with exercises.

3. INTEGRATION MODEL

For this work, we introduce an integration model across user abilities and Learning Analytics (see Fig. 1). Whereas Learning Analytics mostly concerns user interactions with learning activities, we extend to a semantic model that integrates these interactions with user abilities, and is interoperable with Learning Analytics algorithms such as the ELO-rating [5]: as a user interacts with certain learning activities, it can improve its mastery over certain abilities. For example: when a student practices a lot of multiplication exercises, he/she can become better in the "multiplication-ability".

A *User* has – next to its personal data – other data that can be relevant for Learning Analytics, such as having certain *Learning Difficulties* which can be predicted based on user-related data. This User has *Interactions* with *Activities*, which have a certain difficulty determined by e.g. analysis of data from large amount of students. In our case, these interactions are logged with the xAPI. These Activities are related to *User Abilities*, i.e., they train a User ability on a certain difficulty. User Abilities are general abilities that can be further categorized, for example, the "multiplication-ability" is more specific than the "math-ability". The categorization itself is out of scope of this work, but can be modeled using existing standards such as, e.g., the Simple Knowledge Organization System (SKOS) [8]. A User has a certain *Mastery* over these User Abilities. The level of Mastery can be dynamic based on, among others, the Interactions of the User with Activities. This change depends on the item difficulty of the activity, e.g., a user with mastery level 0.4 successfully solving an activity with item difficulty 0.8 has a larger impact than item difficulty 0.1 [5].

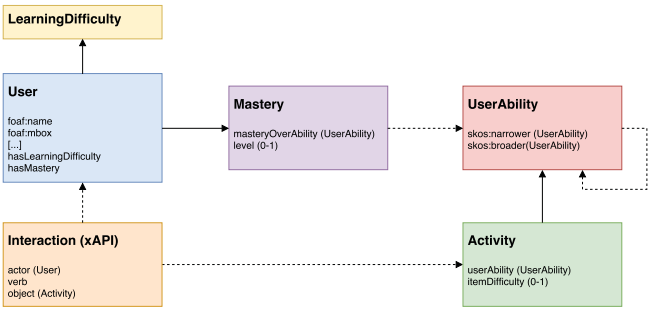


Figure 1: Our proposed integration model across user abilities and Learning Analytics. Available at http://leaps.ilabt.imec.be/specs/

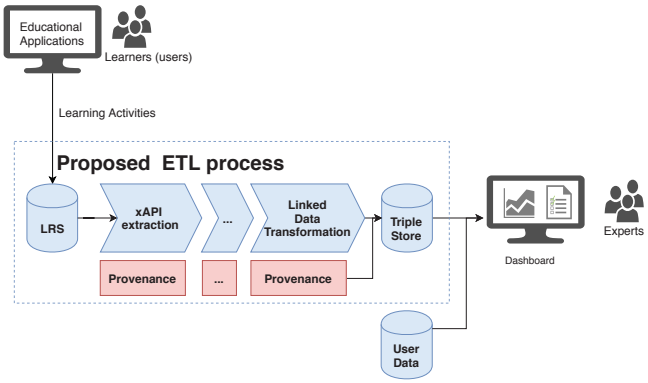


Figure 2: Our proposed ETL pipeline as part of the bigger architecture of the LEAPS project. The Linked Data produced by our proposed pipeline is interlinked with other user data. The xAPI statement extraction component streams newly added statements to one or more transformation components.

4. ETL ARCHITECTURE

Our proposed solution is a layer of Linked Data on top of commonly used Learning Record Systems. Therefore, we utilize an Extract-Transform-Load (ETL) process, where xAPI learning statements are extracted from LRSs, transformed to Linked Data and loaded into a triple store. Provenance of the pipeline's execution is generated and it is also provided as Linked Data.

The starting point of our proposed ETL process are xAPI learning activity statements (see Listing 3 for an example). They are stored in the open-source LRS LearningLocker[5],

[3]TinCan was the former name of xAPI: https://xapi.com/tin-can-experience-api-xapi/

[4]http://edutab.test.iminds.be/specs/serif/

[5]https://learninglocker.net/

from which we extract the data using a streaming approach. Whenever new xAPI statements are inserted to the LRS, we are notified by the underlying storage engine (MongoDB) and start the transformation process. In case of the initial data capture of the LEAPS project, which has already happened, we feed a JSON dump of the learning activities batch-wise to our streaming pipeline.

The transformation process consists of multiple streaming components and therefore new transformations can be plugged in. For the transformation of xAPI statements to Linked Data, we make use of the JSON-LD context of the TinCan2PROV tool introduced by De Nies et al. [4].

Other transformation steps can be added based on the use-case, e.g., if use-case specific xAPI extensions are part of the data. As the case for Anseeuw et al. [2] with their context extension or in other work of us [7], where a privacy usage policy regarding the consented use of the learning data is attached to the xAPI statement as extension.

Additionally to the transformation steps, our proposed ETL process keeps track of provenance, which aims to provide necessary information about the data's origin for later usage. The provenance information per component consists of (i) the name and version of the component; (ii) timestamps of start and end of processing; (iii) used entities (e.g. used JSON-LD contexts, input or output files); (iv) a link to the previous component; and (v) the type of activity (One of the eight sub-classes of the DataActivity class of the GDPRov[6] ontology). The last pipeline step transforms the provenance data to Linked Data, expressing it with the GDPRov[6] vocabulary. At the end of the transformation, the data, as well as collected provenance are available as Linked Data in RDF and are uploaded to a triple store (blazegraph[7]) to be queried with SPARQL. Figure 2 shows our proposed pipeline as part of the bigger LEAPS architecture.

5. THE LEAPS USE CASE

This section demonstrates the interlinking of learning activity data with user-related data, based on the motivating scenario of the ongoing LEAPS project.

Our proposed ETL process transforms learning activity data to Linked Data, with the goal of further data integration. The collected learning activity data are described using the xAPI specification and are using the *SERIF* extension to model the performed exercises. That exercise model allows to describe the difficulty of each exercise.

User-related data describing abilities of users according to the integration model introduced in Section 3 (Listing 1), can be linked with learning activity data describing the exercises with the *SERIF* format (Listing 3). SPARQL queries (such as the one described in Listing 2) can retrieve data which can serve a Learning Analytics algorithm as input.

6. CONCLUSION

One of the goals of Learning Analytics is analysis of diverse types and interrelationships of data. We proposed a provenance-aware ETL process to transform xAPI learning activity statements to Linked Data, with the aim of linking learning data with other user-related data, to assist adaptive learning analytics. We demonstrated how semantic data regarding user abilities can be linked to learning activities. Linked Data allows to combine relevant user data with learning activity data in a seamless way and enable more sophisticated analytics. Combining learning activity data (which contains personal data) with other personal data might cause trouble in the light of data protection [6]. However, the semantic nature of our proposed solution facilitates also privacy-related tasks, as usage policies could be expressed semantically as well [7]. Additionally the provenance produced by our pipeline offers valuable insights in case of an audit.

[6]https://openscience.adaptcentre.ie/ontologies/GDPRov/docs/index-en.html#DataActivity
[7]https://www.blazegraph.com/

```
@prefix foaf: <http://xmlns.com/foaf/0.1/>.
@prefix leaps: <http://edutab.test.iminds.be/specs/datamodel/ontology.owl#>.

:ben foaf:name "Ben De Meester" ;
  foaf:mbox "mailto:ben.demeester@ugent.be"@en;
  [...] ; # more personal data
  leaps:hasLearningDifficulty :dyslexiaA ;
  leaps:hasMastery :benMastery .

:dyslexiaA a leaps:LearningDifficulty .

:benMastery a leaps:Mastery ;
  leaps:masteryOverAbility :skillA ;
  leaps:level "0.6"^^xsd:number .

:skillA a leaps:userAbility ;
  skos:broader :math .
```

Listing 1: Linked Data in turtle format, describing abilities of a user regarding *skillA*.

```
PREFIX : <http://example.com/>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX xapi: <http://semweb.mmlab.be/ns/tincan2prov/>
PREFIX lom: <http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#>
PREFIX leaps: <http://edutab.test.iminds.be/specs/datamodel/ontology.owl#>

SELECT ?name ?masteryLevel ?educationalDifficulty
WHERE {

# Get learning activity user by email address
?agent a xapi:Agent ;
foaf:mbox "mailto:ben.demeester@ugent.be"@en .

?mastery a leaps:Mastery ;
leaps:masteryOverAbility :skillA ;
leaps:level ?masteryLevel .

# Get user data by email address
?user foaf:mbox "mailto:ben.demeester@ugent.be"@en ;
foaf:name ?name ;
leaps:hasMastery ?mastery .

# Get educational difficulty of all learning activities
?xAPIStatement xapi:actor ?agent ;
xapi:object ?object .

?object xapi:definition ?objectDefinition .
?objectDefinition xapi:extensions ?extension .
?extension <http://edutab.test.iminds.be/specs/serif> ?serifExtension .
?serifExtension lom:educationalDifficulty ?educationalDifficulty .
}
```

Listing 2: A SPARQL query, linking user ability data from Listing 1 and generated Linked Data from Listing 4

7. ACKNOWLEDGMENTS

```
{ "timestamp": "2017-01-03T12:26:37.450Z",
  "actor": {
    "objectType": "Agent",
    "mbox": "mailto:ben.demeester@ugent.be" },
  "verb": {
    "id": "http://adlnet.gov/expapi/verbs/attempted" },
  "object": {
    "objectType": "Activity",
    "id": "http://example.com/LezerGame/1/1.0/a/1-2-3-4-5",
    "definition": {
      "name": {
        "nl-BE": "LezerGame 1-1-0-a, deel '1-2-3-4-5'" },
      "type": "http://adlnet.gov/expapi/activities/assessment",
      "extensions": {
        "http://edutab.test.iminds.be/specs/serif": {
          "type": "AssessmentItem",
          "lom:educationalDifficulty": "0.8",
          "itemBody": {
            "value": "Kies welk koekje juist is." },
          "item": [
            { "type": "ChoiceInteraction",
              "properties": [
                { "key": "maxChoices",
                  "value": 1  } ],
              "options": [
                { "label": "bes",
                  "value": "bes" },
                { "label": "bal",
                  "value": "bal"
                } ] } ] } } },
  "context": {
    "contextActivities": {
      "parent": [
        { "id": "http://example.com/LezerGame/1/1.0/a",
          "objectType": "Activity"  } ],
      "category": [
        { "id": "http://id.tincanapi.com/recipe/assessment/general/1"
        } ] } } }
```

Listing 3: A xAPI learning activity statement using the SERIF extension. The statement represents one multiple-choice exercise of the LEAPS project.

```
@prefix xapi: <http://semweb.mmlab.be/ns/tincan2prov/>.
@prefix prov: <http://www.w3.org/ns/prov#>.
@prefix foaf: <http://xmlns.com/foaf/0.1/>.
@prefix leaps: <http://edutab.test.iminds.be/specs/datamodel/ontology.owl#>.
@prefix lom: <http://data.opendiscoveryspace.eu/lom_ontology_ods.owl#>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.


# Linked Data describing the whole statement
_:b12_b0
    a xapi:Statement;
    xapi:actor _:b12_b1;
    xapi:object <http://example.com/LezerGame/1/1.0/a/1-2-3-4-5>.

# Actor part of xAPI statement
_:b12_b1
    a xapi:Agent;
    foaf:mbox "mailto:ben.demeester@ugent.be"@en.

# Object part of xAPI statement
<http://example.com/LezerGame/1/1.0/a/1-2-3-4-5>
    a xapi:Activity;
    xapi:definition _:b12_b6.

# Definition part of the xAPI object
_:b12_b6
    xapi:name "LezerGame 1-1-0-a, deel '1-2-3-4-5'"@nl-be;
    xapi:extensions _:b12_b7.

# The SERIF extension part of the statement
_:b12_b7
    <http://edutab.test.iminds.be/specs/serif> _:b12_b8.

_:b12_b8
    a leaps:AssessmentItem;
    lom:educationalDifficulty "0.8"@en;
    leaps:item _:b12_b9;
    leaps:itemBody _:b12_b14.

_:b12_b9 a leaps:ChoiceInteraction;
    leaps:option _:b12_b12, _:b12_b11, _:b12_b10;
    leaps:properties _:b12_b13.
```

Listing 4: The result of Linked Data transformation from the xAPI statement from Listing 3 (removed unnecessary triples for readability). The JSON-LD contexts of De Nies et al. [4] and De Meester et al. [3] were used to generate Linked Data.

The described research activities were funded by Ghent University, imec, Flanders Innovation & Entrepreneurship (AIO), and the European Union, in the context of the project *LEAPS*. Ruben Verborgh is a postdoctoral fellow of the Research Foundation – Flanders. We also would like to thank the reviewers, whose comments helped to improve this work.

References

1. ISO/IEC 20748-1:2016. Information technology for learning, education and training – Learning analytics interoperability – Part 1: Reference model, 2016-12.

2. Jonas Anseeuw, Stijn Verstichel, Femke Ongenae, Ruben Lagatie, Sylvie Venant, and Filip De Turck. An ontology-enabled context-aware learning record store compatible with the experience api. In *Proceedings of the 8th IC3K Conference, Porto, Portugal*, pages 88–95, November 2016.

3. Ben De Meester, Hajar Ghaem Sigarchian, Tom De Nies, Ruben Verborgh, Frank Salliau, Erik Mannens, and Rik Van de Walle. SERIF: A Semantic ExeRcise Interchange Format. In *Proceedings of the 1st International Workshop on LINKed EDucation*, October 2015.

4. Tom De Nies, Frank Salliau, Ruben Verborgh, Erik Mannens, and Rik Van de Walle. TinCan2PROV: exposing interoperable provenance of learning processes through experience API logs. In *Proceedings of the Linked Learning Workshop*, pages 689–694, May 2015. URL http://www.www2015.it/documents/proceedings/companion/p689.pdf.

5. Arpad E. Elo. *The Rating of Chess Players, Past and Present*. Ishi Press, May 2008. ISBN 978-0923891275.

6. European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, L119:1–88, May 2016. URL http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC.

7. Sven Lieber, Ben De Meester, Anastasia Dimou, and Ruben Verborgh. Privacy law compliance for learning analytics through provenance. In *International Provenance and Annotation Workshop*, 2018. Currently under review.

8. Alistair Miles and Sean Bechhofer. SKOS Simple Knowledge Organization System reference. W3c recommendation, W3C, August 2009. URL https://www.w3.org/TR/skos-reference/.

9. Thomas Rabelo, Manuel Lama, Ricardo R. Amorim, and Juan C. Vidal. Smartlak: A big data architecture for supporting learning analytics services. In *Frontiers in Education 2015*, pages 1–5, Piscataway, NJ, 2015. IEEE. ISBN 978-1-4799-8454-1. .

Detecting, Understanding and Supporting Everyday Learning in Web Search

Ran Yu
L3S Research Center
Hannover, Germany
yu@l3s.de

Ujwal Gadiraju
L3S Research Center
Hannover, Germany
gadiraju@l3s.de

Stefan Dietze
L3S Research Center
Hannover, Germany
dietze@l3s.de

ABSTRACT

Web search is among the most ubiquitous online activities, commonly used to acquire new knowledge and to satisfy learning-related objectives through informational search sessions. The importance of learning as an outcome of web search has been recognized widely, leading to a variety of research at the intersection of information retrieval, human computer interaction and learning-oriented sciences. Given the lack of explicit information, understanding of users and their learning needs has to be derived from their search behavior and resource interactions. In this paper, we introduce the involved research challenges and survey related work on the detection of learning needs, understanding of users, e.g. with respect to their knowledge state, learning tasks and learning progress throughout a search session as well as the actual consideration of learning needs throughout the retrieval and ranking process. In addition, we summarise our own research contributing to the aforementioned tasks and describe our research agenda in this context.

KEYWORDS

Search As Learning, User Modeling, Search Intent Detection, Learning in Web Search

ACM Reference Format:
Ran Yu, Ujwal Gadiraju, and Stefan Dietze. 2018. Detecting, Understanding and Supporting * Everyday Learning in Web Search. In *Proceedings of* . ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Web search is among the most frequent online activities and has become a ubiquitous task. As is common search practice, a coherent search *session*, involving a particular search intent, usually involves several queries as well as one or more breaks in between (cf. [17]).

In particular, *informational* search sessions [7], i.e. sessions pertaining to the search for a particular piece of information expected to be available on the Web, are common and involve a particular learning intent, that is, the intent to acquire knowledge with respect to a certain topic.

, ,

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Whereas platforms dedicated to online learning, such as MOOC environments, are tailored towards improving the learning performance and experience of online users, contemporary search engines have to satisfy a range of use cases, which may or may not involve learning. *Transactional* search sessions [7] are a common example of non-learning related online search. In contrast to actual learning-oriented environments in the online or offline sphere, where certain knowledge about the learning intent, the user as well as the learning task usually is available, such information is lacking in general online search settings. Consequently, heterogeneous features observable throughout a Web search session have to be utilised to derive insights about the learning intent, the user and the actual learning task.

Recently, a range of research works have approached this problem, often summarised unter the '*search as learning (SAL)*' umbrella and involving distinct disciplines such as information retrieval, human computer interaction or machine learning.

This paper attempts to provide an overview of a SAL research agenda by (i) summarising research challenges involved in this context, (ii) discussing related works in the area, (iii) presenting insights into early results of the authors' own work as well as (iv) introducing gaps and future work in this area.

Figure 1 summarises the key emerging research challenges which at the same time define the structure for the remaining sections. *Detecting Learning in Web Search* (Section 2), refers to the process of distinguishing learning-related activities from other, non-learning, activities in general Web search scenarios. *Understanding Users, Learning Tasks, Resources* (Section 3) refers to the challenges involved in inferring information about a user, such as his/her knowledge state, the learning task, such as its complexity, or the involved resources from unstructured behavioral data observable throughout an online search session. Finally, *Supporting Learning through Retrieval and Ranking* (Section 4) refers to the actual consideration of inferred learning needs as part of the retrieval and ranking process or through adapting search interfaces to the user's learning intent.

2 DETECTING LEARNING IN WEB SEARCH

Whereas only a certain amount of Web search sessions include a particular learning need, identifying such sessions becomes a prerequisite to facilitate further applications for understanding and supporting learning.

An established taxonomy from Broder [7] that has been widely used in the Web search context distinguishes between transactional, navigational and informational search sessions, where in particular the latter involve a learning goal, i.e. the intent to acquire knowledge about a particular topic. Specifically, *transactional*

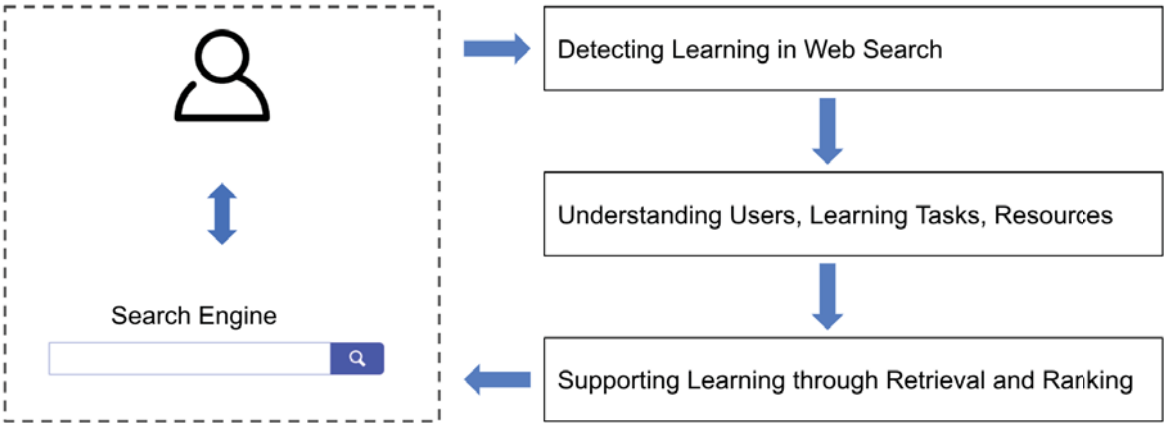


Figure 1: The detecting, understanding, supporting everday learning in Web search pipeline.

search sessions usually aim at conducting a specific online transaction, such as, purchasing a ticket, *navigational* queries merely are aimed at leading the user to a dedicated website. In contrast, *informational* sessions imply the intent of a user to acquire some information assumed to be present on one or more web pages. In this context, the same query, for instance, *Elbphilharmonie* may be used to either buy tickets in a particular concert venue (transactional), to reach the Website https://www.elbphilharmonie.de (*navigational*) or to acquire knowledge about the *Elbphilharmonie* (informational). We inspected a real-world query log, which consists of 913 search sessions, we found that 49.7% of them were informational search sessions with specific learning intent.

By adopting Broderś taxonomy, the task of detecting learning in Web search can be formalised as identifying informational Web search sessions. Here a Web search *session* refers to the search activities within a certain time period that share the same information need. Previous work [17] on segmenting such search sessions have achieved promising performance. Here we do not go into details about the session segmentation task but focus on the automated detection of the intent of Web search.

2.1 Related Works

The classification of Web search queries has been explored in several different scenarios, for instance, to classify a query into one of the categories [23, 24, 27, 29] or subcategories [21] of Broderś taxonomy, or into other classes that are tailored towards specific applications [5, 19, 26].

Early studies on intent classification relied on manual approaches, for instance, by asking users through surveys [7] or by manual annotation of intents through judges [32]. However, while this process does not scale well to large datasets, automated classification approaches have been explored. Both supervised [5, 19, 21, 23, 26, 27, 29] and unsupervised [5, 24] approaches have been applied on the classification of Web search queries. The features utilised in the aforementioned approaches are extracted from query terms [19, 21, 24, 26], user-click behaviors [26, 27, 29], anchor-links [27, 29], Web documents' content [21] and page views [24, 26].

The aforementioned works focus on the classification of single query sessions, often limited to data collected through lab studies. However, recent studies have shown that users information seeking tasks have grown more sophisticated [22] and often require one or more queries across multiple search sessions [1, 25, 28].

2.2 Detecting the Intent of Web Search Sessions from User Interaction Features

In contrast to such previous works, focused on query-based intent detection, our ongoing work to address this problem focuses on automatically detecting the intent of search activities at the level of search sessions.

Approach. We approach the problem of detecting informational Web search sessions with supervised models for classification. We extract 22 features according to multiple dimensions of a search session, structured into three categories, namely features related to *Query* (i.e.features related to number of query terms and the between query similarity), *Session* (i.e. total number of queries issued, session duration related and session breaks related features) and *Browsing* behaviour (i.e. features related to number of clicks, revisited pages and similarity between query and the clicked URL). For the classification model, we have experimented with several different approaches. Considering the scale of the data as well as the number and characteristics of the features, we have opted for Decision Tree (DT), Logistic Regression (LR), Support Vector Machine (SVM) [31] and Random Forest (RF) [6] as classification models. We tune the hyper parameters of each classifier through grid search. The preliminary result of the performance of each classifier is reported below.

Preliminary Result.

We apply our model to a dataset of real-world query logs, which contains 6860 queries from 124 users corresponding to 913 sessions. Each session has been manually annotated by at least two annotators and assigned to one of the three classes. The annotated dataset is available online[1]. The results of using standard precision, recall and F1 score for each individual class, as well as their average

[1]http://l3s.de/~yu/mission_classification

Table 1: Performance of different classifiers.

	Navigational			Informational			Transactional			Weighted average			All
Method	**P**	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**	**Accu**
DT	0.764	0.731	0.747	0.644	0.839	0.728	0.241	**0.076**	**0.116**	0.599	0.653	0.611	0.653
SVM	0.786	**0.760**	**0.773**	**0.656**	0.927	0.768	**0.800**	0.022	0.042	**0.724**	**0.694**	0.623	**0.694**
LR	**0.809**	0.709	0.756	0.651	**0.938**	**0.769**	0.556	0.054	0.099	0.680	0.691	**0.630**	0.691
RF	0.782	0.731	0.756	0.648	0.923	0.761	0.556	0.027	0.052	0.670	0.685	0.617	0.685

across classes and the overall accuracy of the tested classifiers are shown in Table 1. For all configurations, the classification accuracy is above 0.653 and the F1 score is above 0.611, which indicates that the set of features we extracted from user search activities can provide meaningful evidence for detecting the search intent.

We also analyzed the information gain of the selected features, and found that features in the browsing category appear more important than features in other categories, with 2 browsing features ranking at the top 2 positions. Query features are also shown to be effective with 3 features among the top 6. Session-based features have the least contribution among all 3 categories.

For simplicity, we use the term "session" in the remaining of this paper to refer to informational Web search sessions in particular, i.e. sessions which involve a particular learning intent.

2.3 Future Work

The overall classification performance indicates reasonable results on average, in particular transactional sessions appear ambiguous for both human annotators as well as supervised models. For this reason, results indicate that more specific classification tasks are likely to yield superior performance. For instance, an application-specific classifier aimed at targeted advertising may focus on only transactional or informational sessions (depending on the advertised offering), so that binary classification can be applied through a more tailored model. Further more, we found limitations arise in particular from the nature of the experimental dataset and the lack of publicly available, up-to-date query logs, future work will be concerned in particular with the application of similar approaches on a more recent and larger scale dataset. This would enable supervised models which are better reflecting contemporary search behavior and at the same time, utilise a wider variety of features.

3 UNDERSTANDING THE LEARNING PROCESS AND OUTCOMES

Since a sound understanding of learning throughout the search process is required in order to support learning, in this section we summarize the existing efforts in relevant topics and introduce our ongoing works on this task.

There are several factors that potentially affect the learning performance and the required support throughout the search session. These can be classified into three main classes, namely, *user*, e.g. the initial knowledge state and behavioural pattern, *learning task*, such as the task difficulty and novelty, and *resource*, e.g. the complexity or relevance of a resource. These factors are strongly inter-dependent. For instance, the task difficulty is subjective to user's knowledge state, whereas the user's knowledge state is a decisive factor on the resource selection.

Given the sparsity and heterogeneity of data throughout a search session, the works discussed in this section aim at inferring the aforementioned notions by considering a range of features observable throughout a search session.

3.1 Related Works

Previous works assessed the relation between learning and user search behavior from several different perspectives. Eickhoff et al. [11] investigated the correlation between features extracted from search session as well as search engine result page (SERP) documents with learning needs related to either procedural or declarative knowledge. The influence of distinct query types on knowledge gain was studied by Collins-Thompson et al. [9], finding that intrinsically diverse queries lead to increased knowledge gain. Hagen et al. [18] investigated the relation between the writing behavior and the exploratory search pattern of writers and revealed that query terms can be learned while searching and reading. Vakkari [34] provided a structured survey of features indicating learning needs as well as user knowledge and knowledge gain throughout the search process. Zhuang et al. [39] investigated the possibility of using 37 user search behavioral features to predict the user engagement, which correlates with learning, with supervised classifiers. By matching the learning tasks into different learning stages of Anderson and Krathwohl's taxonomy [2], Jansen et al. studied the correlation between search behaviors of 72 participants and their learning stage [20]. Gwizdka et al. [15] proposed to assess learning outcomes in search environments by correlating individual search behaviors with corresponding eye-tracking measures. White et al. [35] investigated the difference between the behavior of domain experts and non-experts in seeking information on the same topic. By analyzing the activity log of experts and non-experts across different domains, the authors found that the distribution of features such as number of queries and query length differed across the levels of expertise. Zhang et al. [37, 38] explored using search behavior as an indicator for the domain knowledge of a user based on data acquired through a lab study ($n = 35$). Further, Cole et al. [8], observed that behavioral patterns provide reliable indicators about the domain knowledge of a user, even if the actual content or topics of queries and documents are disregarded entirely. Gwizdka and Spence [16] have shown that a searcher's perception of task difficulty is a subjective factor that depends on the domain knowledge and some other individual traits. Arguello [3] proposed to use logistic regression to predict task difficulty in a search environment using behavioural features.

The aforementioned prior works have either studied a limited set of features or have addressed only specific learning scenarios and learning types. In particular, the generalizability of knowledge gain measures in previous works has not been investigated.

3.2 Analyzing Knowledge Gain of Users in Informational Search Sessions on the Web

We extend the current understanding of user knowledge gain in informational search sessions. Using real world information needs and search sessions on the Web, we investigate the possibility of using search activity related features to predict knowledge gain (Section 3.3).

In particular, our recent work [14] investigated the impact of information needs on the search behavior and knowledge gain of users. To further the current understanding of the impact of informational search on a user's knowledge, we recruited 500 distinct users from a crowdsourcing platform and orchestrated search sessions spanning 10 different information needs. We followed the recommended guidelines for effective crowdsourcing [12, 13]. By employing scientifically formulated knowledge tests to calibrate a user's knowledge before a search session, and assess it after the session, we were able to quantify knowledge gain. The collected data has been released for the purpose of supporting research in the field[2].

Our investigation revealed a significant effect of information need on user queries and navigational patterns, but no direct effect on the knowledge gain. Users on average exhibited a higher knowledge gain through search sessions pertaining to topics they were less familiar with. For more details and findings please refer to the original paper [14].

3.3 Towards Predicting User Knowledge Gain in Informational Search Sessions

Based on the advanced understanding of the relation between user knowledge and their search behavior, we investigate the possibility of using search activity related features to predict knowledge gain and the knowledge state of a user – avoiding the need for explicit post-search knowledge assessments [36].

In this work [36], we aim at classifying the knowledge state (gain) of a user at the end of a given search session. For the sake of this work, a user's knowledge state with respect to a particular information need is defined as the predicted user capability (accuracy) to correctly respond to a set of test questions about the respective information need. We classify the user knowledge state into 3 classes according to the user capability: low knowledge state, moderate knowledge state and high knowledge state. We hence define the user's knowledge gain as the amount of knowledge state change, and consequently classify the knowledge gain into 3 classes: low knowledge gain, moderate knowledge gain and high knowledge gain.

Approach. We approach the problem with supervised models for classification. To this end, each session is represented by a feature vector, consisting out of 79 features related to: i) query (e.g. number of query terms, query complexity), ii) SERP (e.g. number of clicks, click-through ratio), iii) browsing behaviour (e.g. number of pages viewed, average time stay per page), and iv) mouse movement (e.g. total scroll distance, number of mouseovers). We applied several feature selection techniques on the considered set of feature, and a range of standard models for the classification tasks, namely, Naive Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), Random Forrest (RF) and Multilayer Perceptron (MP).

Preliminary Result. Using the search activity log and the knowledge test data we collected through crowdsourcing (see Section 3.2), we trained and evaluated our classification models. The experimental results underline that a user's knowledge gain and knowledge state can be modeled based on a user's online interactions observable throughout the search process. Through feature analysis, we provide evidence for an improved understanding between individual user behavior and the corresponding knowledge state and change.

3.4 Future Work

As part of future work, we aim to reproduce and refine the findings in more varied search sessions, where durations and learning intents are more diverse, also involving considerably longer/shorter search sessions and, for instance, procedural knowledge rather than intents focused on declarative knowledge only. This would provide the opportunity to observe evolution-oriented features, for instance, considering the evolution of queries, their length and complexity. In addition, in crowd-based quasi experiments understanding of the actual users is very limited and data collected as such is expected to exhibit a certain amount of noise. For these reasons, we aim at conducting equivalent experiments in more controlled lab environments, where reliable information about both user interactions as well as the actual users can be obtained. Furthermore, whereas our previous work has focused on user interaction features, ongoing research investigates resource-centric features which take into account the characteristics of resources involved within the user interactions.

Potential applications for this work include the consideration of user knowledge and its expected learning progress as part of Web search engines and information retrieval approaches, or within informal learning-oriented search settings, such as libraries or knowledge- and resource-centric online platforms.

4 SUPPORTING LEARNING THROUGHOUT WEB SEARCH

The application-oriented objective are concerned with eventually supporting users in their learning tasks through i) optimizing user interaction and interfaces, and ii) enhancing the retrieval and ranking process. In this section, we review the status of existing techniques, and discuss the potential future directions.

4.1 Related Works

User interface and interaction. Learning oriented online platforms (e.g. coursera[3], mooc[4], Didactalia[5]) have been constantly

[2]http://l3s.de/ yu/knowledge_in_search/

[3]https://www.coursera.org/

[4]http://mooc.org/

[5]https://didactalia.net

optimized to improve the learning performance of users. Examples are, for instance, the use of learning dashboards to inform users about his/her learning progress or provide discussion forums to enable collaboration among users. However, within general-purpose search engines, there is a lack of attention for the support of learning, also due to the general-purpose nature of such environments and the variety of tasks conducted there. A central question for research in that area is whether and how interfaces can be adopted to improve learning performance even in such general-purpose environments. An attempt has been made by Arora et al. [4], by aiming at improving user engagement in learning oriented search tasks through providing richer representation of retrieved Web documents. Specifically, they explored methods of finding useful semantic concepts within retrieved documents, with the objective of creating improved document surrogates for presentation in the SERP.

Retrieval and ranking. As current search engines are optimized by considering an information need disregarding the learning intent behind a query, relatively little research has been carried out on optimising retrieval and ranking algorithms towards particular learning needs. For instance, Dave et al. [10] discussed the potential of two ranking models with varied objectives (i.e. paragraph retrieval model, dependency based re-ranking) on enhancing the performance of learning-centric search engines. Recently, Syed and Collins-Thompson [33] proposed to optimize the learning outcome of the vocabulary learning task by selecting a set of documents while considering keyword density and domain knowledge of the learner. Their theoretical framework provides a sound basis for furthering the study on learning-oriented retrieval techniques.

4.2 Future Work

User interface and interaction. Within general-purpose search engines, there is a lack of attention for the support of learning, also due to the general-purpose nature of such environments and the variety of tasks conducted there. A central question for research in that area is whether and how interfaces can be adopted to improve learning performance even in such general-purpose environments. Studies reveal that people engage more in many search tasks involving collaboration with others rather than while searching by themselves [30]. To further this investigation and develop tools to support learning by enabling collaboration between users, our ongoing work is concerned with developing a search interface that encourages experienced learners to guide learners who will use the system in the future and assess its impact on the knowledge gain throughout a search session. Suggestions are ranked according to the feedbacks from experienced learners. Throughout large-scale quasi-experiments and facilitated by pre- and post-tests, we aim to quantify the influence of the collaborative search interface on the learning outcome. Future work is concerned with alternative means to improve interfaces and interactions towards increasing the learning outcomes during Web search.

Retrieval and ranking. In Section 3.3, we have discussed means to infer a user's knowledge state (gain) in online search sessions. On this basis, future work is aimed at optimizing ranking algorithms to recommend resources that fit not only the traditional notion of an information need, but also a user's knowledge state. Whereas traditional ranking algorithms tend to suggest Web documents disregarding, for instance, a user's reading level, improved ranking techniques will favor resources which are neither too easy nor too hard for a particular user's need. This builds on the assumption that, based on the assessment of the relation between a Web resource and user's knowledge gain, a ranking algorithm can recommend resources not only fitting into the user's knowledge state, but also maximizing the user's knowledge gain and learning efficiency.

5 CONCLUSIONS

This paper has provided an overview of challenges and research approaches towards detecting, understanding and supporting learning throughout Web search. One crucial challenge in this context is the lack of explicit information about users, their learning intent, task or progress throughout an online search session, requiring the utilisation of a wide variety of informal features to derive such information. In particular, supervised machine learning techniques and extensive feature analysis have been deployed as part of previous work, yet works are usually focused on specific learning scenarios, isolated feature sets or single-query scenarios, rather than entire search sessions.

Another major obstacle is the lack of large-scale datasets to facilitate SAL research by providing both diverse features of user interactions and behavior as well as high-quality ground truth data about the involved users, their knowledge state and knowledge gain throughout the captured search sessions.

In addition to summarising research challenges and related works, we have introduced some of our own contributions to the respective tasks. These consist of supervised approaches for detecting learning-related (informational) search sessions, for predicting the knowledge state and gain of online users and the preliminary analysis of experimentally obtained search sessions and the correlation of observed variables with user knowledge. Ongoing and future work will expand on these works, consider more varied feature sets, in particular resource-centric features, and will in particular be concerned with obtaining, providing and analysing search session data collected in more controlled lab environments.

REFERENCES

[1] Eugene Agichtein, Ryen W. White, Susan T. Dumais, and Paul N. Bennet. 2012. Search, Interrupted: Understanding and Predicting Search Task Continuation. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '12)*. ACM, New York, NY, USA, 315–324.

[2] Lorin W Anderson, David R Krathwohl, P Airasian, K Cruikshank, R Mayer, P Pintrich, J Raths, and M Wittrock. 2001. A taxonomy for learning, teaching and assessing: A revision of BloomâĂŹs taxonomy. *New York. Longman Publishing. Artz, AF, & Armour-Thomas, E.(1992). Development of a cognitive-metacognitive framework for protocol analysis of mathematical problem solving in small groups. Cognition and Instruction* 9, 2 (2001), 137–175.

[3] Jaime Arguello. 2014. Predicting Search Task Difficulty.. In *ECIR*, Vol. 14. 88–99.

[4] Piyush Arora. 2015. Promoting User Engagement and Learning in Amorphous Search Tasks. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 1051–1051.

[5] Ricardo Baeza-Yates, Liliana Calderón-Benavides, and Cristina González-Caro. 2006. The Intention Behind Web Queries. In *Proceedings of the 13th International Conference on String Processing and Information Retrieval (SPIRE'06)*. Springer-Verlag, Berlin, Heidelberg, 98–109.

[6] Leo Breiman. 2001. Random Forests. *Mach. Learn.* 45, 1 (Oct. 2001), 5–32.

[7] Andrei Broder. 2002. A Taxonomy of Web Search. *SIGIR Forum* 36, 2 (9 2002), 3–10. Issue Fall 2002.

[8] Michael J Cole, Jacek Gwizdka, Chang Liu, Nicholas J Belkin, and Xiangmin Zhang. 2013. Inferring user knowledge level from eye movement patterns. *Information Processing & Management* 49, 5 (2013), 1075–1091.

[9] Kevyn Collins-Thompson, Soo Young Rieh, Carl C Haynes, and Rohail Syed. 2016. Assessing learning outcomes in web search: A comparison of tasks and query strategies. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval*. ACM, 163–172.

[10] Kushal Dave, Vasudeva Varma, et al. 2014. Computational advertising: Techniques for targeting relevant ads. *Foundations and Trends® in Information Retrieval* 8, 4–5 (2014), 263–418.

[11] Carsten Eickhoff, Jaime Teevan, Ryen White, and Susan Dumais. 2014. Lessons from the journey: a query log analysis of within-session learning. In *Proceedings of the 7th ACM international conference on Web search and data mining*. ACM, 223–232.

[12] Ujwal Gadiraju, Ricardo Kawase, Stefan Dietze, and Gianluca Demartini. 2015. Understanding malicious behavior in crowdsourcing platforms: The case of online surveys. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 1631–1640.

[13] Ujwal Gadiraju, Jie Yang, and Alessandro Bozzon. 2017. Clarity is a Worthwhile Quality–On the Role of Task Clarity in Microtask Crowdsourcing. In *Proceedings of the 28th ACM Conference on Hypertext and Social Media*. ACM, 5–14.

[14] Ujwal Gadiraju, Ran Yu, Stefan Dietze, and Peter Holtz. 2018. Analyzing Knowledge Gain of Users in Informational Search Sessions on the Web. In *2018 ACM on Conference on Human Information Interaction and Retrieval (CHIIR)*. ACM.

[15] Jacek Gwizdka and Xueshu Chen. 2016. Towards Observable Indicators of Learning on Search.. In *SAL@ SIGIR*.

[16] Jacek Gwizdka and Ian Spence. 2006. What can searching behavior tell us about the difficulty of information tasks? A study of Web navigation. *Proceedings of the Association for Information Science and Technology* 43, 1 (2006), 1–22.

[17] Matthias Hagen, Jakob Gomoll, Anna Beyer, and Benno Stein. 2013. From Search Session Detection to Search Mission Detection. In *10th International Conference Open Research Areas in Information Retrieval (OAIR 13)*, John P. McDermott (Ed.). ACM, 85–92.

[18] Matthias Hagen, Martin Potthast, Michael Völske, Jakob Gomoll, and Benno Stein. 2016. How Writers Search: Analyzing the Search and Writing Logs of Non-fictional Essays. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval*. ACM, 193–202.

[19] Jian Hu, Gang Wang, Fred Lochovsky, Jian-tao Sun, and Zheng Chen. 2009. Understanding User's Query Intent with Wikipedia. In *Proceedings of the 18th International Conference on World Wide Web (WWW '09)*. ACM, New York, NY, USA, 471–480.

[20] Bernard J Jansen, Danielle Booth, and Brian Smith. 2009. Using the taxonomy of cognitive learning to model online searching. *Information Processing & Management* 45, 6 (2009), 643–663.

[21] Bernard J. Jansen, Danielle L. Booth, and Amanda Spink. 2008. Determining the Informational, Navigational, and Transactional Intent of Web Queries. *Inf. Process. Manage.* 44, 3 (May 2008), 1251–1266.

[22] Rosie Jones and Kristina Lisa Klinkner. 2008. Beyond the session timeout: automatic hierarchical segmentation of search topics in query logs. In *Proceedings of the 17th ACM conference on Information and knowledge management*. ACM, 699–708.

[23] In-Ho Kang and GilChang Kim. 2003. Query Type Classification for Web Document Retrieval. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval (SIGIR '03)*. ACM, New York, NY, USA, 64–71.

[24] A. Kathuria, B. J. Jansen, C. Hafernik, and A. Spink. 2010. Classifying the user intent of web queries using k-means clustering. *Internet Research* 20, 5 (2010), 563–581.

[25] Alexander Kotov, Paul N. Bennett, Ryen W. White, Susan T. Dumais, and Jaime Teevan. 2011. Modeling and Analysis of Cross-session Search Tasks. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '11)*. ACM, New York, NY, USA, 5–14.

[26] Elad Kravi, Ido Guy, Avihai Mejer, David Carmel, Yoelle Maarek, Dan Pelleg, and Gilad Tsur. 2016. One Query, Many Clicks: Analysis of Queries with Multiple Clicks by the Same User. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. ACM, 1423–1432.

[27] Uichin Lee, Zhenyu Liu, and Junghoo Cho. 2005. Automatic Identification of User Goals in Web Search. In *Proceedings of the 14th International Conference on World Wide Web (WWW '05)*. ACM, New York, NY, USA, 391–400.

[28] Jingjing Liu and Nicholas J. Belkin. 2010. Personalizing Information Retrieval for Multi-session Tasks: The Roles of Task Stage and Task Type. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '10)*. ACM, New York, NY, USA, 26–33.

[29] Yiqun Liu, Min Zhang, Liyun Ru, and Shaoping Ma. 2006. Automatic Query Type Identification Based on Click Through Information. In *Proceedings of the Third Asia Conference on Information Retrieval Technology (AIRS'06)*. Springer-Verlag, Berlin, Heidelberg, 593–600.

[30] Meredith Ringel Morris. 2007. Collaborating alone and together: Investigating persistent and multi-user web search activities. In *Proceedings of international ACM SIGIR conference on research and development in information retrieval (SIGIR 2007)*. 23–27.

[31] John C. Platt. 1999. Advances in Kernel Methods. MIT Press, Cambridge, MA, USA, Chapter Fast Training of Support Vector Machines Using Sequential Minimal Optimization, 185–208.

[32] Daniel E. Rose and Danny Levinson. 2004. Understanding User Goals in Web Search. In *Proceedings of the 13th International Conference on World Wide Web*. ACM, 13–19.

[33] Rohail Syed and Kevyn Collins-Thompson. 2017. Retrieval algorithms optimized for human learning. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 555–564.

[34] Pertti Vakkari. 2016. Searching as learning: A systematization based on literature. *Journal of Information Science* 42, 1 (2016), 7–18.

[35] Ryen W White, Susan T Dumais, and Jaime Teevan. 2009. Characterizing the influence of domain expertise on web search behavior. In *Proceedings of the second ACM international conference on web search and data mining*. ACM, 132–141.

[36] Ran Yu, Ujwal Gadiraju, Peter Holtz, Markus Rokicki, Philipp Kemkes, and Stefan Dietze. 2018. Predicting User Knowledge Gain in Informational Search Sessions. In *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM.

[37] Xiangmin Zhang, Michael Cole, and Nicholas Belkin. 2011. Predicting users' domain knowledge from search behaviors. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. ACM, 1225–1226.

[38] Xiangmin Zhang, Jingjing Liu, Michael Cole, and Nicholas Belkin. 2015. Predicting users' domain knowledge in information retrieval using multiple regression analysis of search behaviors. *Journal of the Association for Information Science and Technology* 66, 5 (2015), 980–1000.

[39] Mengdie Zhuang, Gianluca Demartini, and Elaine G Toms. 2017. Understanding engagement through search behaviour. In *International Conference on Information and Knowledge Management, Proceedings*. Association for Computing Machinery, 1957–1966.